

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 11:23:43 ; Search time 3628.41 Seconds  
(without alignments)  
78.116 Million cell updates/sec

Title: US-09-913-325-4  
Perfect score: 21  
Sequence: 1 cagcagcagagcttcacatc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: EST:\*  
2: em\_estba:\*  
3: em\_esthum:\*  
4: em\_estin:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	21	100.0	85	9	AM276802 xp66a01.x
2	21	100.0	100	10	BF920141 MR1-NT017
3	21	100.0	102	10	BF958934 PM1-NN120
4	21	100.0	119	9	AW901233
5	21	100.0	142	10	DA5267
6	21	100.0	148	10	BF846357
7	21	100.0	149	10	R47195
8	21	100.0	169	10	BE766895
9	21	100.0	177	9	AL048592
10	21	100.0	183	10	B1032792
11	21	100.0	184	10	BF958930 PM1-NN120
12	21	100.0	185	10	B1036860
13	21	100.0	196	9	A1745406
14	21	100.0	197	10	B1036862
15	21	100.0	201	10	BF957666
16	21	100.0	203	10	BF957653
17	21	100.0	207	10	BF957858

c 18	21	100.0	224	9	AM161224	AM161224 au70a10.y
c 19	21	100.0	226	10	BF935119	BF935119 MR4-NT014
c 20	21	100.0	236	10	BF948789	BF948789 MR3-NN021
c 21	21	100.0	238	10	BF923639	BF923639 MR4-NT014
c 22	21	100.0	239	10	BG88974	BG88974 HOA21-I-C
c 23	21	100.0	243	9	AA336628	AA336628 EST41242
c 24	21	100.0	244	10	BE766870	BE766870 RC2-NT011
c 25	21	100.0	248	10	BF923643	BF923643 MR4-NT014
c 26	21	100.0	249	10	BF945175	BF945175 PM1-NN120
c 27	21	100.0	252	10	BF923633	BF923633 MR4-NT014
c 28	21	100.0	252	10	BF963107	BF963107 PM1-NN120
c 29	21	100.0	253	10	BF887875	BF887875 QV2-TN017
c 30	21	100.0	253	10	B1041998	B1041998 MR4-NT014
c 31	21	100.0	256	10	B1041248	B1041248 MR4-NT014
c 32	21	100.0	258	10	BF961150	BF961150 PM1-NN120
c 33	21	100.0	262	10	BF947155	BF947155 MR3-NN021
c 34	21	100.0	263	10	BE899032	BE899032 601682590
c 35	21	100.0	265	10	BF961152	BF961152 PM1-NN120
c 36	21	100.0	268	10	BF920149	BF920149 MR1-NT017
c 37	21	100.0	269	10	BF948799	BF948799 MR3-NN021
c 38	21	100.0	272	10	BF956582	BF956582 PM1-NN120
c 39	21	100.0	273	10	BF947794	BF947794 MR3-NN021
c 40	21	100.0	273	10	BF956584	BF956584 PM1-NN120
c 41	21	100.0	274	10	B1036852	B1036852 MR4-NT014
c 42	21	100.0	283	10	BF947214	BF947214 MR4-NN021
c 43	21	100.0	287	10	B1041483	B1041483 MR4-NT014
c 44	21	100.0	293	9	AA318247	AA318247 EST20305
c 45	21	100.0	297	10	BF949666	BF949666 QV2-NN119

## ALIGNMENTS

RESULT 1  
AM276802/c  
LOCUS xp66a01.x1 NCI-CGAP\_Ov39 Homo sapiens CDNA clone IMAGE:2745288 3'  
DEFINITION similar to gb:xl4723 CLUSTERIN PRECURSOR (HUMAN);, mRNA sequence.  
ACCESSION AM276802  
VERSION AM276802.1 GI:6663832  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 85)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
Tumor Gene Index  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-rt@mail.nih.gov](mailto:cgapbs-rt@mail.nih.gov)  
Tissue Procurement: Andrew Berchuck M.D., John Gillespie M.D.,  
Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Kitzman, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILN at:  
[www.bio.litl.gov/btrp/image/image.html](http://www.bio.litl.gov/btrp/image/image.html)

## FEATURES

source  
1. 85  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2745288"  
/clone\_lib="NCI-CGAP\_Ov39"  
/sex="female"  
/tissue\_type="papillary serous ovarian metastasis"  
/lab\_host="DH10B"  
/note="Organ: ovary; Vector: pAMP10; CDNA made by oligo-dT  
Possible reversed clone: polyT not found  
Seq primer: -40up from Gibco.  
Location/Qualifiers

priming. Non-directionally cloned into the UDG sites of  
 PAMP10. Size-selected on agarose gel, average insert  
 size 500 bp. Primary library: non-amplified. cDNA  
 Library Preparation: David B. Kitzman, Ph.D (NCI).  
 Reference: Kitzman et al. (1996) Cancer Research  
 56:5380-5383."

BASE COUNT 18 a 21 c 30 g 16 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcacatcat 21  
 |||||||||||||||||||  
 Db 82 CAGCAGCAGAGCTTCATCAT 62

RESULT 2  
 LOCUS BF920141 100 bp mRNA linear EST 19-JAN-2001  
 DEFINITION MRI-NT0179-071100-003-f02 NT0179 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF920141  
 VERSION BF920141.1 GI:12316029  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 100)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE  
 COMMENT Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR1&t2=MR1-NT0179-  
 071100-003-f02&t3=2000-11-07&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 11  
 High quality sequence stop: 100.  
 Location/Qualifiers

FEATURES  
 source  
 1..100  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NT0179"  
 /dev\_stage="Adult"

/note="Organ: nervous\_tumor; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 21 a 34 c 20 g 25 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 100;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcacatcat 21  
 |||||||||||||||||||  
 Db 56 CAGCAGCAGAGCTTCATCAT 76

RESULT 3  
 LOCUS BF958934/c 102 bp mRNA linear EST 22-JAN-2001  
 DEFINITION PMI-NN1200-011200-009-b07 NN1200 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF958934  
 VERSION BF958934.1 GI:12376209  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 102)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE  
 COMMENT Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1&t2=PM1-NN1200-  
 011200-009-b07&t3=2000-12-01&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 11  
 High quality sequence stop: 101.  
 Location/Qualifiers

FEATURES  
 source  
 1..102  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NN1200"  
 /dev\_stage="Adult"

/note="Organ: nervous\_normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 21 a 25 c 38 g 18 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcacatcat 21  
 |||||||||||||||||||  
 Db 91 CAGCAGCAGAGCTTCATCAT 71

RESULT 4

AW901233/C  
LOCUS AW901233 119 bp mRNA linear EST 24-MAY-2000  
DEFINITION CM4-NN1011-100300-110-g04 NN1011 Homo sapiens cDNA, mRNA sequence.  
ACCESSION AW901233  
VERSION AW901233.1 GI:8065542  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 119)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
JOURNAL Contact: Simpson A.J.G.  
MEDLINE Laboratory of Cancer Genetics  
COMMENT Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4-NN1011-100  
300-110-g04&ts=2000-03-10&tl=1)  
Seq primer: puc 18 forward  
High quality sequence start: 11  
High quality sequence stop: 118.  
Location/Qualifiers  
1..119  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN1011"  
/dev\_stage="Adult"  
/note="Organ: nervous, normal. Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT 21 a 28 c 46 g 23 t 1 others  
ORIGIN  
Query Match 100.0%; Score 21; DB 9; Length 119;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 cagcagcagcttcacat 21  
|||||  
Db 103 CAGCAGCAGCTTCATCAT 83  
RESULT 5  
LOCUS D45267 142 bp mRNA linear EST 30-DEC-1995  
DEFINITION HUMHGL194 Human cerebral cortex Homo sapiens cDNA, mRNA sequence.  
ACCESSION D45267  
VERSION D45267.1 GI:1136645  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 142)  
AUTHORS Takahashi, N., Hashida, H., Zhao, N., Misumi, Y. and Sakaki, Y.  
TITLE High-density cDNA filter analysis of the expression profiles of the  
genes preferentially expressed in human brain  
JOURNAL Gene 164, 219-227 (1995)  
MEDLINE 96069586  
COMMENT Contact: Nobuaki Takahashi  
Institute of Medical Science  
University of Tokyo  
Shirokanedai 4-6-1, Minato-ku, Tokyo, Japan 108  
Tel: 03-5449-5625  
Fax: 03-5449-5445.  
Location/Qualifiers  
1..142  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Human cerebral cortex"  
/note="Adult male cerebral cortex tissue."  
BASE COUNT 26 a 31 c 55 g 30 t  
ORIGIN  
Query Match 100.0%; Score 21; DB 10; Length 142;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 cagcagcagcttcacat 21  
|||||  
Db 79 CAGCAGCAGCTTCATCAT 59  
RESULT 6  
LOCUS BF846357 148 bp mRNA linear EST 16-JAN-2001  
DEFINITION PM1-EN0065-231000-002-b01 EN0065 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF846357  
VERSION BF846357.1 GI:12233611  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 148)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
JOURNAL Contact: Simpson A.J.G.  
MEDLINE Laboratory of Cancer Genetics  
COMMENT Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1-EN0065-  
231000-002-b01&ts=2000-10-23&tl=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 110.  
Location/Qualifiers  
1..148  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="EN0065"  
/dev\_stage="Adult"

/note="Organ: lung\_normal; Vector: puc18; Site\_1: Sma1; Site\_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 24 a 56 c 41 g 27 t

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 148;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cagcagcagagcttcacatcat 21  
|||||  
Db 34 CAGCAGCAGAGCTTCATCAT 54

RESULT 7  
R47195 149 bp mRNA linear EST 12-DEC-1995  
LOCUS R47195  
DEFINITION CBS-389 Subtractive cDNA library ocular ciliary body Homo sapiens  
cDNA clone CBS-389 5' end similar to TRPM-2 (clusterin) (accession number M64722), mRNA sequence.  
R47195  
VERSION R47195.1 GI:807537  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 149)  
AUTHORS Escribano,J., Ortego,J. and Coca-Prados,M.  
TITLE Isolation and characterization of cell-specific cDNA clones from a subtractive library of the ocular ciliary body of a single normal human donor: transcription and synthesis of plasma proteins

JOURNAL J. Blochem. (Tokyo) 118 (5), 921-931 (1995)  
MEDLINE 96318503  
COMMENT Contact: Coca-Prados, M.  
Department of Ophthalmology and Visual Science  
Yale University Medical School  
330 Cedar Street, New Haven, CT 06520-8061  
Tel: 203/7852742  
Fax: 203/7856123  
Email: miguel.coca-prados@quickmail.yale.edu  
Seq primer: T3.

FEATURES  
source  
Location/Qualifiers  
1..149  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CBS-389"  
/clone\_lib="Subtractive cDNA library ocular ciliary body"  
/note="Vector: pBluescript II SK; Site\_1: EcoRI; Site\_2: XhoI; A subtractive cDNA library was developed by hybridizing antisense, single-stranded phagemid DNA (ssDNA) (as pBluescript SK-) from the ocular ciliary body cDNA library (target) of a 34-year-old female donor in lambda-uni-zap XR with biotinylated sense RNA of an ocular cell line cDNA library (driver) in the same vector."

BASE COUNT 33 a 31 c 56 g 29 t

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 149;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cagcagcagagcttcacatcat 21  
|||||  
Db 71 CAGCAGCAGAGCTTCATCAT 51

RESULT 8  
BE766895 169 bp mRNA linear EST 19-SEP-2000  
LOCUS BE766895  
DEFINITION RC2-W0110-050600-013-f03 NT0110 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE766895  
VERSION BE766895.1 GI:10196819  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 169)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Mitsuikuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=RC2-W0110-050600-013-f03&tl3=2000-06-05&td=1)

Seq primer: puc 18 forward  
High quality sequence start: 18  
High quality sequence stop: 169.  
Location/Qualifiers  
1..169  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NT0110"  
/dev\_stage="Adult"  
/note="Organ: nervous\_tumor; Vector: puc18; Site\_1: Sma1; Site\_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 37 a 36 c 59 g 37 t

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 169;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cagcagcagagcttcacatcat 21  
|||||  
Db 59 CAGCAGCAGAGCTTCATCAT 39

RESULT 9  
AL048592 177 bp mRNA linear EST 01-MAR-2000  
LOCUS AL048592  
DEFINITION DKF2P586H092\_r1 586 (synonym: hute1) Homo sapiens cDNA clone  
ACCESSION AL048592  
VERSION AL048592.1 GI:4729143  
KEYWORDS EST.  
SOURCE human.



ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 177)  
AUTHORS Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
TITLE EST (Poustka, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Poustka A.J.  
Department Lehrach  
Max-Planck-Institute for Molecular Genetics  
Innestrasse 73, 14195 Berlin, Germany  
Tel: +49-30-84131623  
Fax: +49-30-84131128  
Email: poustka@mping-berlin-dahlem.mpg.de  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No sl sequence available.  
This clone (DKFZp586H092) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
FEATURES  
source  
1. 177  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="DKFZp586H092"  
/clone\_1lb="586 (synonym: hute1)"  
/tissue\_type="uterus"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI/MluI"  
BASE COUNT 40 a 41 c 64 g 32 t  
ORIGIN  
Query Match 100.0%; Score 21; DB 9; Length 177;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 cagcagcagagcttcacat 21  
|||||  
Db 77 CACCACGACAGCTTCATCAT 57  
RESULT 10 183 bp mRNA linear EST 14-JUN-2001  
BI032792  
LOCUS BI032792  
DEFINITION M4-NN0188-220101-204-e01 NN0188 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BI032792  
VERSION BI032792.1 GI:14439418  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 183)  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=MR4&t2=MR4-NN0188-  
220101-204-e01&t3=2001-01-22&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 183.  
FEATURES  
source  
1. 183  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN0188"  
/dev\_stage="Adult"  
/note="organ: nervous normal; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT 28 a 72 c 48 g 35 t  
ORIGIN  
Query Match 100.0%; Score 21; DB 10; Length 183;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 cagcagcagagcttcacat 21  
|||||  
Db 69 CAGCAGCAGAGCTTCATCAT 89  
RESULT 11 184 bp mRNA linear EST 22-JAN-2001  
BF958930  
LOCUS BF958930  
DEFINITION PM1-NN1200-011200-009-g08 NN1200 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF958930  
VERSION BF958930.1 GI:12376205  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 184)  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=PM1&t2=PM1-NN1200-  
011200-009-g08&t3=2000-12-01&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 7  
High quality sequence stop: 184.

```
FEATURES
source
    Location/Qualifiers
    1.184
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="NN1200"
    /dev_stage="Adult"
    /note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;
    Site_2: SmaI; A mini-library was made by cloning products
    derived from ORESTES PCR (U.S. Letters Patent application
    No. 196,716 - Ludwig Institute for Cancer Research)
    profiles into the puc 18 vector. Reverse transcription of
    tissue mRNA and cDNA amplification were performed under
    low stringency conditions."
BASE COUNT      47 a      36 c      61 g      40 t
ORIGIN

Query Match      100.0%; Score 21; DB 10; Length 184;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 cagcagcagagctctcatcat 21
        |||
Db      46 cagcagcagagctctcatcat 26

RESULT 12
BI036860/c      185 bp      mRNA      linear      EST 14-JUN-2001
DEFINITION      MR4-NT0140-080101-209-c09 NT0140 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BI036860
VERSION      BI036860.1 GI:14443486
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 185)
    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
    Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
    Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
    Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
    ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
    Simpson,A.J.
    Shotgun sequencing of the human transcriptome with ORF expressed
    sequence tags
    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL
COMMENT
    Contact: Simpson A.J.G.
    Laboratory of Cancer Genetics
    Ludwig Institute for Cancer Research
    Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
    Brazil
    Tel: +55-11-2704922
    Fax: +55-11-2707001
    Email: asimpson@ludwig.org.br
    This sequence was derived from the FAPESP/LICR Human Cancer Genome
    Project. This entry can be seen in the following URL
    (http://www.ludwig.org.br/scripts/gethtml2.pl?cli=MR4&cl2=MR4-NT0140-
    080101-209-c09&cl3=2001-01-08&cl4=1)
    Seq primer: puc 18 forward
    High quality sequence start: 17
    High quality sequence stop: 134.
FEATURES
source
    Location/Qualifiers
    1.185
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="NT0140"
    /dev_stage="Adult"
    /note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI;
    Site_2: SmaI; A mini-library was made by cloning products
    derived from ORESTES PCR (U.S. Letters Patent application
    No. 196,716 - Ludwig Institute for Cancer Research)
    profiles into the puc 18 vector. Reverse transcription of
    tissue mRNA and cDNA amplification were performed under
    low stringency conditions."
BASE COUNT      41 a      42 c      62 g      40 t
ORIGIN

Query Match      100.0%; Score 21; DB 10; Length 185;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 cagcagcagagctctcatcat 21
        |||
Db      43 cagcagcagagctctcatcat 23

RESULT 13
AI745406
LOCUS      AI745406      196 bp      mRNA      linear      EST 17-DEC-1999
DEFINITION      wc37401.x1 NCI-CGAP Pr28 Homo sapiens cDNA clone IMAGE:2320801 3'
    similar to gb:U14723 CLUSTERIN PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION      AI745406
VERSION      AI745406.1 GI:5113694
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 196)
    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
    Tumor Gene Index
    Unpublished (1997)
    Contact: Robert Strausberg, Ph.D.
    Email: cgaps@email.nih.gov
    Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
    Emmert-Buck, M.D., Ph.D.
    cDNA Library Preparation: M. Bento Soares, Ph.D.
    DNA Library Arrayed by: Greg Lennon, Ph.D.
    DNA sequencing by: Washington University Genome Sequencing Center
    Clone distribution: NCI-CGAP clone distribution information can be
    found through the I.M.A.G.E. Consortium/LLNL at:
    www.bio.llnl.gov/birp/image/image.html
    Insert Length: 264 Std Error: 0.00
    Seq primer: -40UP from Gibco.
FEATURES
source
    Location/Qualifiers
    1.196
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="IMAGE:2320801"
    /clone_lib="NCI-CGAP_Pr28"
    /sex="male"
    /dev_stage="adult"
    /lab_host="DH10B"
    /note="Organ: prostate; Vector: p7T73D-Pac (Pharmacia)
    with a modified polylinker; Plasmid DNA from the
    normalized library NCI-CGAP Pr22 was prepared, and ss
    circles were made in vitro. Following HAP purification,
    this DNA was used as tracer in a subtractive hybridization
    reaction. The driver was PCR-amplified cDNAs from a pool
    of 5,000 clones made from the same library (clonoids
    985608-986759, 1101192-1101959, and 1217928-1220615)."
    Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      39 a      60 c      35 g      62 t
ORIGIN

Query Match      100.0%; Score 21; DB 9; Length 196;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 cagcagcagagctctcatcat 21
        |||
```

Db 167 CAGCAGCAGAGCTTCATCAT 187

RESULT 14  
BI036862/c 197 bp mRNA linear EST 14-JUN-2001

LOCUS BI036862  
DEFINITION MR4-NT0140-080101-209-d10 NT0140 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BI036862  
VERSION BI036862.1 GI:14443488  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 197)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.G.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

AUTHORS  
CONTACT: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL.  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR4&tc=MR4-NT0140-080101-209-d10&ts=2001-01-08&tl=1)  
Seq primer: puc 18 forward  
High quality sequence start: 27  
High quality sequence stop: 197.  
Location/Qualifiers  
1. 197  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NT0140"  
/dev\_stage="Adult"  
/note="Organ: nervous\_tumor; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 49 a 40 c 66 g 42 t

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 197;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cagcagcagagcttcacat 21  
|||||

Db 55 CAGCAGCAGAGCTTCATCAT 35

RESULT 15  
BF957666/c 201 bp mRNA linear EST 22-JAN-2001

LOCUS BF957666  
DEFINITION PM1-NN1200-251100-007-c09 NN1200 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF957666  
VERSION BF957666.1 GI:12374941  
KEYWORDS EST.

SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 201)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.G.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

AUTHORS  
CONTACT: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL.  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1&tc=PM1-NN1200-251100-007-c09&ts=2000-11-25&tl=1)  
Seq primer: puc 18 forward  
High quality sequence start: 14  
High quality sequence stop: 201.  
Location/Qualifiers  
1. 201  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NN1200"  
/dev\_stage="Adult"  
/note="Organ: nervous\_normal; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 52 a 39 c 39 t

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 201;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cagcagcagagcttcacat 21  
|||||

Db 59 CAGCAGCAGAGCTTCATCAT 39

Search completed: May 17, 2002, 15:25:36  
Job time: 14513 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 15:25:46 : Search time 147.64 Seconds  
(without alignments)  
66.583 Million cell updates/sec

Title: US-09-913-325-4  
Perfect score: 21  
Sequence: 1 cagcagcagagcttcattcatc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 424068 segs, 234053524 residues

Total number of hits satisfying chosen parameters: 848136

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_New.\*  
1: /cgn2\_6/ptodata/1/pna/PCT\_NEW\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	21	100.0	1614	6	US-10-119-428-31 Sequence 31, App1
C 2	21	100.0	1839	7	US-60-365-384-115 Sequence 115, App
C 3	16.4	78.1	3380	5	US-09-930-213-252 Sequence 252, App
C 4	16.2	77.1	32	5	US-09-784-403A-6 Sequence 6, App11
C 5	16.2	77.1	139	6	US-10-126-704-12 Sequence 12, App1
C 6	16.2	77.1	235	5	US-09-540-210B-23471 Sequence 23471, A
C 7	16.2	77.1	236	5	US-09-540-210B-31214 Sequence 31214, A
C 8	16.2	77.1	257	5	US-09-975-254-17477 Sequence 17477, A
C 9	16.2	77.1	292	5	US-09-539-331D-3257 Sequence 3257, Ap
C 10	16.2	77.1	319	6	US-10-106-698-2234 Sequence 2234, Ap
C 11	16.2	77.1	324	6	US-10-099-926-918 Sequence 918, App
C 12	16.2	77.1	325	6	US-10-097-105-650 Sequence 650, App
C 13	16.2	77.1	330	6	US-10-097-105-664 Sequence 664, App
C 14	16.2	77.1	409	6	US-10-097-105-264 Sequence 264, App
C 15	16.2	77.1	412	5	US-09-954-531-790 Sequence 790, App
C 16	16.2	77.1	434	6	US-10-097-105-595 Sequence 595, App
C 17	16.2	77.1	471	5	US-09-991-150-393 Sequence 95, App
C 18	16.2	77.1	471	5	US-10-119-480-95 Sequence 95, App1
C 19	16.2	77.1	488	5	US-09-602-148A-26 Sequence 26, App1
C 20	16.2	77.1	544	5	US-09-664-249B-1 Sequence 1, App11
C 21	16.2	77.1	742	6	US-10-006-063A-335 Sequence 335, App
C 22	16.2	77.1	742	6	US-10-006-117A-335 Sequence 335, App
C 23	16.2	77.1	742	6	US-10-006-130A-335 Sequence 335, App
C 24	16.2	77.1	742	6	US-10-006-172A-335 Sequence 335, App
C 25	16.2	77.1	742	6	US-10-006-768A-335 Sequence 335, App
C 26	16.2	77.1	742	6	US-10-017-527A-335 Sequence 335, App

C 27	16.2	77.1	742	6	US-10-017-610A-335 Sequence 335, App
C 28	16.2	77.1	742	6	US-10-006-041A-335 Sequence 335, App
C 29	16.2	77.1	742	6	US-10-006-818A-335 Sequence 335, App
C 30	16.2	77.1	742	6	US-10-012-121A-335 Sequence 335, App
C 31	16.2	77.1	742	6	US-10-015-386A-335 Sequence 335, App
C 32	16.2	77.1	742	6	US-10-015-387A-335 Sequence 335, App
C 33	16.2	77.1	742	6	US-10-121-062-359 Sequence 335, App
C 34	16.2	77.1	742	6	US-10-013-907A-335 Sequence 335, App
C 35	16.2	77.1	742	6	US-10-015-499A-335 Sequence 335, App
C 36	16.2	77.1	742	6	US-10-017-867A-335 Sequence 335, App
C 37	16.2	77.1	742	6	US-10-017-867A-335 Sequence 335, App
C 38	16.2	77.1	973	7	US-60-365-384-399 Sequence 399, App
C 39	16.2	77.1	6002	6	US-10-126-704-4 Sequence 4, App11
C 40	16.2	77.1	162450	6	US-10-126-704-1 Sequence 1, App11
C 41	15.8	75.2	1146	5	US-09-540-209B-3032 Sequence 3032, Ap
C 42	15.8	75.2	2680	5	US-09-053-375B-651 Sequence 651, App
C 43	15.8	75.2	2999	6	US-10-106-698-2003 Sequence 2003, Ap
C 44	15.8	75.2	6921	5	US-09-053-375B-93 Sequence 93, App1
C 45	15.8	75.2	6921	6	US-10-117-982-117 Sequence 117, App

## ALIGNMENTS

RESULT 1  
US-10-119-428-31/c  
Sequence 31, Application US/10119428  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Xu, Chongjun  
APPLICANT: Wehman, Tom  
APPLICANT: Ren, Feiyang  
APPLICANT: Ma, Yunging  
APPLICANT: Zhou, Ping  
APPLICANT: Zhao, Qing A.  
APPLICANT: Yang, Yonghong  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: Novel Nucleic Acids and  
FILE REFERENCE: 789CIP2  
CURRENT APPLICATION NUMBER: US/10/119,428  
CURRENT FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: 09/596,193  
PRIOR FILING DATE: 2000-06-17  
PRIOR APPLICATION NUMBER: 09/574,454  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: 09/519,705  
PRIOR FILING DATE: 2000-03-07  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: pt\_FL-genes Version 1.0  
SEQ ID NO 31  
LENGTH: 1614  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (116)..(1465)  
US-10-119-428-31  
Query Match 100.0% Score 21; DB 6; Length 1614;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 cagcagcagagcttcattcatc 21  
Db 136 cagcagcagagcttcattcatc 116  
RESULT 2  
US-60-365-384-115/c

```
; Sequence 115, Application US/60365384
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-ru
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Weinman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ghosh, Malabika
; APPLICANT: Zhao, Qing A.
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aiding J.
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrul
; APPLICANT: Wang, Gezhi
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 814
; CURRENT APPLICATION NUMBER: US/60/365,384
; CURRENT FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 666
; SOFTWARE: PE_FL-genes Version 6.0
; SEQ ID NO 115
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (132)..(1316)
; US-60-365-384-115
```

```
Query Match          100.0%; Score 21; DB 7; Length 1839;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 cagcagcagagctctcatcat 21  
|||||

DB 152 CAGCAGCAGAGCTTTCATCAT 132

```
RESULT 3
; Sequence 252, Application US/0930213
; GENERAL INFORMATION:
; APPLICANT: ROSENTHAL, ANDRE
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHAFER, REINHARD
; APPLICANT: ZUBER, JOHANNES
; APPLICANT: TCHE-NITSE, OLEG
; APPLICANT: GRIPS, MARTIN
; APPLICANT: HELLMINGEL, MARTIN
; APPLICANT: SCHMITZ, ANNE-CHANTAL
; APPLICANT: SERS, CHRISTINE
; TITLE OF INVENTION: DETECTION OF DIFFERENTIAL GENE EXPRESSIONS
; FILE REFERENCE: ALBRE-14
; CURRENT APPLICATION NUMBER: US/09/930,213
; CURRENT FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: DE 10004102.7
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 885
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 252
; LENGTH: 3380
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-930-213-252
```

```
Query Match          78.1%; Score 16.4; DB 5; Length 3380;
Best Local Similarity 94.4%; Pred. No. 73;
```

```
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 agcagcagagctctcatc 19
|
DB 190 AACAGCAGAGCTTTCATC 173
```

```
RESULT 4
; Sequence 5, Application US/09784403A
; GENERAL INFORMATION:
; APPLICANT: Barbour, Eric
; APPLICANT: Eucilaire Meyer, Terry
; APPLICANT: Eid Saad, Mohammed
; TITLE OF INVENTION: Novel Maize Promoters
; FILE REFERENCE: 35718/208067
; CURRENT APPLICATION NUMBER: US/09/784,403A
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: US 60/107,201
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: US 60/103,294
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 09/410,935
; PRIOR FILING DATE: 1999-10-04
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 32
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Gene specific primer 1 for Gcs-2
; US-09-784-403A-6
```

```
Query Match          77.1%; Score 16.2; DB 5; Length 32;
Best Local Similarity 85.7%; Pred. No. 58;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

QY 1 cagcagcagagctctcatcat 21  
|||||

DB 3 cagcagcagagctctcatcat 23

```
RESULT 5
; Sequence 12, Application US/10126704
; GENERAL INFORMATION:
; APPLICANT: Bouguetelret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-
; FILE REFERENCE: 44.US.DIV
; CURRENT APPLICATION NUMBER: US/10/126,704
; CURRENT FILING DATE: 2002-04-20
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 12
; LENGTH: 139
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-126-704-12
```

```
Query Match          77.1%; Score 16.2; DB 6; Length 139;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

QY 1 cagcagcagagctctcatcat 21  
|||||

Db 3 cagaggaagagcttcacatc 23

RESULT 6  
US-09-540-210B-23471  
; Sequence 23471, Application US/09540210B  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
; FILE REFERENCE: PD-1037 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,210B  
; CURRENT FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: 08/972,899  
; PRIOR FILING DATE: November 18, 1997  
; PRIOR APPLICATION NUMBER: 08/395,244  
; PRIOR FILING DATE: February 27, 1995  
; PRIOR APPLICATION NUMBER: 08/722,922  
; PRIOR FILING DATE: September 27, 1996  
; PRIOR APPLICATION NUMBER: 60/005,526  
; PRIOR FILING DATE: September 29, 1995  
; PRIOR APPLICATION NUMBER: 08/824,029  
; PRIOR FILING DATE: March 25, 1997  
; PRIOR APPLICATION NUMBER: 60/014,010  
; PRIOR FILING DATE: March 25, 1996  
; PRIOR APPLICATION NUMBER: 08/826,847  
; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996  
; PRIOR APPLICATION NUMBER: 08/903,555  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/862,178  
; PRIOR FILING DATE: May 22, 1997  
; PRIOR APPLICATION NUMBER: 60/018,217  
; PRIOR FILING DATE: May 23, 1996  
; PRIOR APPLICATION NUMBER: 08/881,589  
; PRIOR FILING DATE: June 24, 1997  
; PRIOR APPLICATION NUMBER: 60/021,275  
; PRIOR FILING DATE: June 25, 1996  
; PRIOR APPLICATION NUMBER: 08/903,802  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/023,308  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/905,881  
; PRIOR FILING DATE: August 1, 1997  
; PRIOR APPLICATION NUMBER: 60/025,204  
; PRIOR FILING DATE: August 1, 1996  
; PRIOR APPLICATION NUMBER: 08/903,471  
; PRIOR FILING DATE: July 30, 1997  
; PRIOR APPLICATION NUMBER: 60/025,478  
; PRIOR FILING DATE: July 31, 1996  
; PRIOR APPLICATION NUMBER: 08/903,556  
; PRIOR FILING DATE: July 31, 1997  
; PRIOR APPLICATION NUMBER: 60/025,217  
; PRIOR FILING DATE: August 22, 1996  
; PRIOR APPLICATION NUMBER: 08/937,142  
; PRIOR FILING DATE: September 23, 1997  
; PRIOR APPLICATION NUMBER: 60/026,598  
; PRIOR FILING DATE: September 24, 1996  
; PRIOR APPLICATION NUMBER: 08/960,746  
; PRIOR FILING DATE: October 29, 1997  
; PRIOR APPLICATION NUMBER: 60/030,144  
; PRIOR FILING DATE: October 30, 1996  
; PRIOR APPLICATION NUMBER: 08/826,847  
; PRIOR FILING DATE: April 10, 1997  
; PRIOR APPLICATION NUMBER: 60/015,533  
; PRIOR FILING DATE: April 10, 1996

;; PRIOR APPLICATION NUMBER: 08/755,524  
;; PRIOR FILING DATE: November 22, 1996  
;; PRIOR APPLICATION NUMBER: 60/007,495  
;; PRIOR FILING DATE: November 22, 1995  
;; PRIOR APPLICATION NUMBER: 09/021,031  
;; PRIOR FILING DATE: February 10, 1998  
;; PRIOR APPLICATION NUMBER: 60/039,325  
;; PRIOR FILING DATE: February 13, 1997  
;; PRIOR APPLICATION NUMBER: 09/035,172  
;; PRIOR FILING DATE: March 4, 1998  
;; PRIOR APPLICATION NUMBER: 60/040,431  
;; PRIOR FILING DATE: March 5, 1997  
;; PRIOR APPLICATION NUMBER: 09/041,894  
;; PRIOR FILING DATE: March 12, 1998  
;; PRIOR APPLICATION NUMBER: 60/040,199  
;; PRIOR FILING DATE: March 14, 1997  
;; PRIOR APPLICATION NUMBER: 09/050,817  
;; PRIOR FILING DATE: March 30, 1998  
;; PRIOR APPLICATION NUMBER: 60/043,792  
;; PRIOR FILING DATE: April 11, 1997  
;; PRIOR APPLICATION NUMBER: 09/074,999  
;; PRIOR FILING DATE: May 8, 1998  
;; PRIOR APPLICATION NUMBER: 60/048,431  
;; PRIOR FILING DATE: May 29, 1997  
;; PRIOR APPLICATION NUMBER: 09/107,592  
;; PRIOR FILING DATE: June 30, 1998  
;; PRIOR APPLICATION NUMBER: 60/052,751  
;; PRIOR FILING DATE: July 1, 1997  
;; PRIOR APPLICATION NUMBER: 09/094,079  
;; PRIOR FILING DATE: June 9, 1998  
;; PRIOR APPLICATION NUMBER: 60/049,975  
;; PRIOR FILING DATE: June 13, 1997  
;; NUMBER OF SEQ ID NOS: 35654  
;; SOFTWARE: PERL Program  
;; SEQ ID NO: 23471  
;; LENGTH: 235  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc.feature  
;; OTHER INFORMATION: Incyte ID No: hu00253322  
US-09-540-210B-23471

Query Match 77.1%; Score 16.2; DB 5; Length 235;  
Best Local Similarity 85.7%; Pred. No. 71;  
Matches 18; Conservative 0; Mismatches 3; Indels 0;  
Gaps 0;

QY 1 cagcagcagagcttcacatc 21  
Db 160 cagcagcagagcttcacatc 180

RESULT 7  
US-09-540-210B-31214/C  
; Sequence 31214, Application US/09540210B  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
; FILE REFERENCE: PD-1037 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,210B  
; CURRENT FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: 08/972,899  
; PRIOR FILING DATE: November 18, 1997  
; PRIOR APPLICATION NUMBER: 08/395,244  
; PRIOR FILING DATE: February 27, 1995  
; PRIOR APPLICATION NUMBER: 08/722,922  
; PRIOR FILING DATE: September 27, 1996

```

; PRIOR APPLICATION NUMBER: 60/005,526
; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/903,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431

```

```

; PRIOR FILING DATE: May 29, 1997
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO: 31214
; LENGTH: 256
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00338737
US-09-540-210B-31214

```

```

Query Match          77.1%; Score 16.2; DB 5; Length 256;
Best Local Similarity 85.7%; Pred. No. 71;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0.

```

```

Qy      1 cagcagcagagcttcacat 21
Db      64 CTGATCAGAGCTTCATCCT 44

```

```

RESULT      8
US-09-975-254-17477/c
; Sequence 17477, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(15309)B
; CURRENT APPLICATION NUMBER: US/09/975,254
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/263,191
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO: 17477
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 700954386H1
US-09-975-254-17477

```

```

Query Match          77.1%; Score 16.2; DB 5; Length 257;
Best Local Similarity 85.7%; Pred. No. 71;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0.

```

```

Qy      1 cagcagcagagcttcacat 21
Db      253 CAGCAGCAGAGGTTGCATCAT 233

```

```

RESULT      9
US-09-539-331D-3257/c
; Sequence 3257, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE

```



```

; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539,331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 3257
; LENGTH: 292
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00740013
; NAME/KEY: unsure
; LOCATION: 14-15, 54, 113-114
; OTHER INFORMATION: a, t, c, g, or other
; US-09-539-331D-3257

```

```

Query Match          77.1%; Score 16.2; DB 5; Length 292;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 cagcagcagagcttcctcat 21
Db 148 CTGCATCAGAGCTCTTCATCCT 128

```

```

RESULT 10
US-10-106-698-2234/c
; Sequence 2234, Application US/10106698
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: P8005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2234
; LENGTH: 319
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (305)..(305)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-106-698-2234

```

```

Query Match          77.1%; Score 16.2; DB 6; Length 319;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 cagcagcagagcttcctcat 21
Db 102 CATCGAGAGCTCTTCATCAT 82

```

```

RESULT 11
US-10-099-926-918
; Sequence 918, Application US/10099926
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun

```

```

; APPLICANT: Secrist, Heather
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547C2
; CURRENT APPLICATION NUMBER: US/10/099,926
; CURRENT FILING DATE: 2002-03-17
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 918
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-099-926-918

```

```

Query Match          77.1%; Score 16.2; DB 6; Length 324;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 cagcagcagagcttcctcat 21
Db 143 cagcagcagagcttcctcat 163

```

```

RESULT 12
US-10-097-105-650/c
; Sequence 650, Application US/10097105
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.504C1
; CURRENT APPLICATION NUMBER: US/10/097,105
; CURRENT FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 650
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-097-105-650

```

```

Query Match          77.1%; Score 16.2; DB 6; Length 325;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 cagcagcagagcttcctcat 21
Db 183 CACGAGAGAGCTTCATCATCAT 163

```

```

RESULT 13
US-10-097-105-664/c
; Sequence 664, Application US/10097105
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.504C1
; CURRENT APPLICATION NUMBER: US/10/097,105
; CURRENT FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 664
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-105-664

```

```

Query Match          77.1%; Score 16.2; DB 6; Length 330;
Best Local Similarity 85.7%; Pred. No. 73;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 cagcagcagagcttcacatc 21
    ||||||||| | |||||||
Db 184 CACCAGCAGGCATCATCAT 164

```

```

RESULT 14
US-10-097-105-264

```

```

; Sequence 264, Application US/10097105
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504C1
; CURRENT APPLICATION NUMBER: US/10/097,105
; CURRENT FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 264
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 139, 218, 223, 261, 398
; OTHER INFORMATION: n = A,T,C or G
US-10-097-105-264

```

```

Query Match          77.1%; Score 16.2; DB 6; Length 409;
Best Local Similarity 85.7%; Pred. No. 74;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy- 1 cagcagcagagcttcacatc 21
      ||||||||| | |||||||
Db 143 cagcagcagagcttcacatc 163

```

```

RESULT 15
US-09-954-531-790

```

```

; Sequence 790, Application US/09954531
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392

```

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 790
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-790

```

```

Query Match          77.1%; Score 16.2; DB 5; Length 412;
Best Local Similarity 85.7%; Pred. No. 75;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 cagcagcagagcttcacatc 21
    ||||||||| | |||||||
Db 282 caatgcagagcttcacatc 302

```

```

Search completed: May 17, 2002, 17:56:43
Job time: 9177 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 14:24:53 ; Search time 5965.48 Seconds  
(without alignments)  
76.160 Million cell updates/sec

Title: US-09-913-325-4  
Perfect score: 21  
Sequence: 1 cagcagcagcttcacatc 21

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues  
Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/US007\_PCTUS\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/pna/US080\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq:\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq:\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq:\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq:\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq:\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq:\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq:\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq:\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq:\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq:\*  
16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq:\*  
17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:\*  
18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:\*  
19: /cgn2\_6/ptodata/2/pna/US095\_COMB.seq:\*  
20: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq:\*  
21: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq:\*  
22: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq:\*  
23: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq:\*  
24: /cgn2\_6/ptodata/2/pna/US095E\_COMB.seq:\*  
25: /cgn2\_6/ptodata/2/pna/US095F\_COMB.seq:\*  
26: /cgn2\_6/ptodata/2/pna/US095G\_COMB.seq:\*  
27: /cgn2\_6/ptodata/2/pna/US095H\_COMB.seq:\*  
28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq:\*  
29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq:\*  
30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq:\*  
31: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq:\*  
32: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq:\*  
33: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:\*  
34: /cgn2\_6/ptodata/2/pna/US098D\_COMB.seq:\*  
35: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq:\*  
36: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq:\*  
37: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq:\*  
38: /cgn2\_6/ptodata/2/pna/US100\_COMB.seq:\*  
39: /cgn2\_6/ptodata/2/pna/US101\_COMB.seq:\*  
40: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq:\*  
41: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq:\*  
42: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq:\*  
43: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq:\*

SUMMARIES									
Result	Query				Description				
No.	Score	Match	Length	ID					
1	21	100.0	21	1	PCT-US00-04875-4	Sequence 4, Appl1			
2	21	100.0	21	34	US-09-913-325-4	Sequence 4, Appl1			
3	21	100.0	21	35	US-09-944-326-4	Sequence 4, Appl1			
4	21	100.0	21	37	US-10-080-794-4	Sequence 4, Appl1			
5	21	100.0	177	20	US-09-534-840-10325	Sequence 10325, A			
6	21	100.0	177	20	US-09-534-840-10415	Sequence 10415, A			
7	21	100.0	161	8	US-08-485-657-19	Sequence 19, Appl1			
8	21	100.0	181	17	US-09-366-380-19	Sequence 19, Appl1			
9	21	100.0	190	20	US-09-534-840-10784	Sequence 10784, A			
10	21	100.0	192	8	US-09-534-840-9748	Sequence 9748, Ap			
11	21	100.0	195	20	US-08-485-657-14	Sequence 14, Appl1			
12	21	100.0	195	17	US-09-366-380-14	Sequence 14, Appl1			
13	21	100.0	199	20	US-09-534-840-9752	Sequence 9752, Ap			
14	21	100.0	200	14	US-09-021-031-1091	Sequence 1091, Ap			
15	21	100.0	200	20	US-09-534-840-9505	Sequence 9505, Ap			
16	21	100.0	200	42	US-60-039-325-1091	Sequence 1091, Ap			
17	21	100.0	203	20	US-09-534-840-10336	Sequence 10336, A			
18	21	100.0	203	20	US-09-534-840-9736	Sequence 9736, Ap			
19	21	100.0	209	20	US-09-534-840-10340	Sequence 10340, A			
20	21	100.0	210	20	US-09-534-840-10381	Sequence 10381, A			
21	21	100.0	213	13	US-08-918-181-1516	Sequence 1516, Ap			
22	21	100.0	213	20	US-09-534-840-9416	Sequence 9416, Ap			
23	21	100.0	213	41	US-60-023-236-1516	Sequence 1516, Ap			
24	21	100.0	214	20	US-09-534-840-10521	Sequence 10521, A			
25	21	100.0	219	20	US-09-534-840-10313	Sequence 10313, A			
26	21	100.0	222	28	US-09-716-472-2881	Sequence 2881, Ap			
27	21	100.0	223	20	US-09-534-840-10438	Sequence 10438, A			
28	21	100.0	226	13	US-08-961-526-363	Sequence 363, App			
29	21	100.0	226	20	US-09-534-840-9461	Sequence 9461, App			
30	21	100.0	226	20	US-09-534-840-10349	Sequence 10349, A			
31	21	100.0	227	20	US-09-534-840-10374	Sequence 10374, A			

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

```

C 32 21 100.0 228 20 US-09-534-840-10328 Sequence 10328, A
C 33 21 100.0 228 20 US-09-534-840-10330 Sequence 10330, A
C 34 21 100.0 229 20 US-09-534-840-10386 Sequence 10386, A
C 35 21 100.0 231 20 US-09-534-840-10382 Sequence 10382, A
C 36 21 100.0 232 20 US-09-534-840-10342 Sequence 10342, A
C 37 21 100.0 235 20 US-09-534-840-10420 Sequence 10420, A
C 38 21 100.0 235 20 US-09-534-840-10765 Sequence 10765, A
C 39 21 100.0 237 20 US-09-534-840-10442 Sequence 10442, A
C 40 21 100.0 239 20 US-09-534-840-10410 Sequence 10410, A
C 41 21 100.0 240 20 US-09-534-840-10376 Sequence 10376, A
C 42 21 100.0 241 20 US-09-534-840-10351 Sequence 10351, A
C 43 21 100.0 244 20 US-09-534-840-10308 Sequence 10308, A
C 44 21 100.0 245 20 US-09-534-840-10427 Sequence 10427, A
C 45 21 100.0 247 20 US-09-534-840-10332 Sequence 10332, A

```

## ALIGNMENTS

```

RESULT 1
PCT-US00-04875-4
; Sequence 4, Application PC/TUS0004875
; GENERAL INFORMATION:
; APPLICANT: Gleave, Martin
; APPLICANT: Rennie, Paul S.
; APPLICANT: Miyake, Hideaki
; APPLICANT: Nelson, Colleen
; TITLE OF INVENTION: ANTISENSE THERAPY FOR CANCER
; FILE REFERENCE: UBC.P-020-WO
; CURRENT APPLICATION NUMBER: PCT/US00/04875
; CURRENT FILING DATE: 2000-02-25
; EARLIER APPLICATION NUMBER: 60/121,726
; EARLIER FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 21
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: antisense TRPM-2 ODN
PCT-US00-04875-4

```

```

Query Match 100.0%; Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cagcagcagagcttcctcatc 21
DB 1 cagcagcagagcttcctcatc 21

```

```

RESULT 2
US-09-913-325-4
; Sequence 4, Application US/09913325
; GENERAL INFORMATION:
; APPLICANT: Gleave, Martin
; APPLICANT: Rennie, Paul S.
; APPLICANT: Miyake, Hideaki
; APPLICANT: Nelson, Colleen
; TITLE OF INVENTION: TRPM-2 ANTISENSE THERAPY
; FILE REFERENCE: UBC.P-020-WO
; CURRENT APPLICATION NUMBER: US/09/913,325
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/121,726
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 21
; TYPE: DNA
; ORGANISM: HUMAN

```

```

; FEATURE:
; OTHER INFORMATION: antisense TRPM-2 ODN
US-09-913-325-4

```

```

Query Match 100.0%; Score 21; DB 34; Length 21;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cagcagcagagcttcctcatc 21
DB 1 cagcagcagagcttcctcatc 21

```

```

RESULT 3
US-09-944-326-4
; Sequence 4, Application US/09944326
; GENERAL INFORMATION:
; APPLICANT: Gleave, Martin
; APPLICANT: Rennie, Paul S.
; APPLICANT: Miyake, Hideaki
; APPLICANT: Nelson, Colleen
; TITLE OF INVENTION: TRPM-2 ANTISENSE THERAPY
; FILE REFERENCE: UBC.P-020-2
; CURRENT APPLICATION NUMBER: US/09/944,326
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 60/121,726
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 09/913,325
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 21
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: antisense TRPM-2 ODN
US-09-944-326-4

```

```

Query Match 100.0%; Score 21; DB 35; Length 21;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cagcagcagagcttcctcatc 21
DB 1 cagcagcagagcttcctcatc 21

```

```

RESULT 4
US-10-080-794-4
; Sequence 4, Application US/10080794
; GENERAL INFORMATION:
; APPLICANT: Gleave, Martin
; APPLICANT: Rennie, Paul S.
; APPLICANT: Miyake, Hideaki
; APPLICANT: Nelson, Colleen
; APPLICANT: Monia, Brett P.
; TITLE OF INVENTION: TRPM-2 ANTISENSE THERAPY USING AN OLIGONUCLEOTIDE
; FILE REFERENCE: UBC.P-020-3
; CURRENT APPLICATION NUMBER: US/10/080,794
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/121,726
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 09/913,325
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 09/944,326
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4

```

LENGTH: 21  
TYPE: DNA  
ORGANISM: HUMAN  
FEATURE:  
OTHER INFORMATION: antisense TRPM-2 ODN  
US-10-080-794-4

Query Match 100.0%; Score 21; DB 37; Length 21;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cagcagcagagcttcctcatcat 21  
|||||  
DB 1 cagcagcagagcttcctcatcat 21

RESULT 5  
US-09-534-840-10325/c  
Sequence 10325, Application US/09534840  
GENERAL INFORMATION:  
APPLICANT: Seilhamer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MOLEC  
FILE REFERENCE: PD-1010 CIP  
CURRENT APPLICATION NUMBER: US/09/534,840  
CURRENT FILING DATE: 1992-07-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 13675  
SOFTWARE: PERL Program  
SEQ ID NO 10325  
LENGTH: 177  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No: hu01113942  
US-09-534-840-10325

Query Match 100.0%; Score 21; DB 20; Length 177;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cagcagcagagcttcctcatcat 21  
|||||  
DB 88 CAGCAGCAGAGCTTCATCATCAT 68

RESULT 6  
US-09-534-840-10415/c  
Sequence 10415, Application US/09534840  
GENERAL INFORMATION:  
APPLICANT: Seilhamer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MOLEC  
FILE REFERENCE: PD-1010 CIP  
CURRENT APPLICATION NUMBER: US/09/534,840  
CURRENT FILING DATE: 1992-07-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 13675  
SOFTWARE: PERL Program  
SEQ ID NO 10415  
LENGTH: 177  
TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No: hu01055183  
NAME/KEY: unsure  
LOCATION: 17, 40  
OTHER INFORMATION: a, t, c, g, or other  
US-09-534-840-10415

Query Match 100.0%; Score 21; DB 20; Length 177;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cagcagcagagcttcctcatcat 21  
|||||  
DB 88 CAGCAGCAGAGCTTCATCATCAT 68

RESULT 7  
US-08-485-657-19/c  
Sequence 19, Application US/08485657  
GENERAL INFORMATION:  
APPLICANT: Kirschling, Deborah J  
APPLICANT: Gudkov, Andrei  
APPLICANT: Roninson, Igor B  
TITLE OF INVENTION: Genes and Genetic Elements Associated  
TITLE OF INVENTION: with Sensitivity to Cisplatin  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: ALLEGRETTI & WITCOFF, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,657  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/199,900  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Noonan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 93,354-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 181 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-485-657-19

Query Match 100.0%; Score 21; DB 8; Length 181;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cagcagcagagcttcctcatcat 21  
|||||  
DB 73 CAGCAGCAGAGCTTCATCATCAT 53

```
RESULT 8
US-09-366-380-19/c
; Sequence 19, Application US/09366380
; GENERAL INFORMATION:
; APPLICANT: Kirschling, Deborah J
; APPLICANT: Gudkov, Andrei
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Genes and Genetic Elements Associated
; TITLE OF INVENTION: with Sensitivity to Cisplatin
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/366,380
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/199,900
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noonan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-N
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-366-380-19

Query Match          100.0%; Score 21; DB 17; Length 181;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcctcatc 21
    |||
Db 73 CAGCAGCAGAGCTTCATCAT 53

RESULT 9
US-09-534-840-10784/c
; Sequence 14, Application US/09534840
; GENERAL INFORMATION:
; APPLICANT: Kirschling, Deborah J
; APPLICANT: Gudkov, Andrei
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Genes and Genetic Elements Associated
; TITLE OF INVENTION: with Sensitivity to Cisplatin
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; Prior application data removed - consult PALM or file wrapper
```

```
; NUMBER OF SEQ ID NOS: 13675
; SOFTWARE: PERL Program
; SEQ ID NO: 10784
; LENGTH: 190
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: hu01342214
US-09-534-840-10784

Query Match          100.0%; Score 21; DB 20; Length 190;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcctcatc 21
    |||
Db 88 CAGCAGCAGAGCTTCATCAT 68

RESULT 10
US-09-534-840-9748/c
; Sequence 9748, Application US/09534840
; GENERAL INFORMATION:
; APPLICANT: Selthamer, Jeffrey J
; APPLICANT: Deleage, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullaly, Sara J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MO
; FILE REFERENCE: PD-1010 CIP
; CURRENT APPLICATION NUMBER: US/09/534,840
; PRIOR APPLICATION DATA removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 13675
; SOFTWARE: PERL Program
; SEQ ID NO: 9748
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: hu01285157
US-09-534-840-9748

Query Match          100.0%; Score 21; DB 20; Length 192;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcctcatc 21
    |||
Db 157 CAGCAGCAGAGCTTCATCAT 137

RESULT 11
US-08-485-657-14/c
; Sequence 14, Application US/08485657
; GENERAL INFORMATION:
; APPLICANT: Kirschling, Deborah J
; APPLICANT: Gudkov, Andrei
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Genes and Genetic Elements Associated
; TITLE OF INVENTION: with Sensitivity to Cisplatin
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ALLEGRETTI & WITCOFF, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
```

```

;
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,657
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/199,900
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noonan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-485-657-14

Query Match          100.0%; Score 21: DB 8; Length 195;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcctcatc 21
    |||
Db 87 CAGCAGCAGAGCTTCCTCATC 67

RESULT 12
US-09-366-380-14/c
; Sequence 14, Application US/09366380
; GENERAL INFORMATION:
; APPLICANT: Kirschling, Deborah J
; APPLICANT: Gudkov, Andrei
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Genes and Genetic Elements Associated
; TITLE OF INVENTION: with Sensitivity to Cisplatin
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/366,380
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/199,900
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noonan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-N
;
;
```

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
;
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-366-380-14

Query Match          100.0%; Score 21: DB 17; Length 195;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcctcatc 21
    |||
Db 87 CAGCAGCAGAGCTTCCTCATC 67

RESULT 13
US-09-534-840-9752/c
; Sequence 9752, Application US/09534840
; GENERAL INFORMATION:
; APPLICANT: Sellhammer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Muliahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MO
; FILE REFERENCE: PD-1010 CIP
; CURRENT APPLICATION NUMBER: US/09/534,840
; CURRENT FILING DATE: 1992-07-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 13675
; SOFTWARE: PERL Program
; SEQ ID NO 9752
; LENGTH: 199
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: incyte ID No: hu01227727
; NAME/KEY: unsure
; LOCATION: 118
; OTHER INFORMATION: a, t, c, g, or other
; US-09-534-840-9752

Query Match          100.0%; Score 21: DB 20; Length 199;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcctcatc 21
    |||
Db 87 CAGCAGCAGAGCTTCCTCATC 67

RESULT 14
US-09-021-031-1091/c
; Sequence 1091, Application US/09021031
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Naughton, Rebecca E.
;
;
```

APPLICANT: Klingler, Tod M.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
NUMBER OF INVENTION: BLADDER CARCINOMA  
CORRESPONDENCE ADDRESS: 1882  
ADDRESS: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/021,031  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/039,325  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: CERRONE, MICHAEL C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PD-0319P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 853-0535  
TELEFAX: (415) 845-4165  
INFORMATION FOR SEQ ID NO: 1091:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 200 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
CLONE: 2807488H1  
US-09-021-031-1091

Query Match 100.0%; Score 21; DB 14; Length 200;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cagcagcagagtcctcatcat 21  
|||||  
Db 80 CAGCAGCAGACTCTTCATCAT 60

RESULT 15  
US-09-534-840-9505/C  
Sequence 9505, Application US/09534840  
GENERAL INFORMATION:  
APPLICANT: Selihamer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mollahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MOLEC  
FILE REFERENCE: PD-1010 CIP  
CURRENT APPLICATION NUMBER: US/09/534,840  
CURRENT FILING DATE: 1992-07-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 13675  
SOFTWARE: PERL Program  
SEQ ID NO 9505  
LENGTH: 200  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No: hu00709579  
NAME/KEY: unsure  
LOCATION: 86, 142, 156, 186  
OTHER INFORMATION: a, t, c, g, or other  
US-09-534-840-9505

Query Match 100.0%; Score 21; DB 20; Length 200;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 cagcagcagagtcctcatcat 21  
|||||  
Db 80 CAGCAGCAGAGTCTTCATCAT 60

Search completed: May 17, 2002, 17:56:00  
Job time: 12667 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 12:33:46 ; Search time 111.1 Seconds  
(without alignments)  
46.429 Million cell updates/sec

Title: US-09-913-325-4

Perfect score: 21

Sequence: 1 cagcagcagagcttcatacat 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	181	2	US-08-485-657A-19
2	21	100.0	181	5	PCR-US95-02303-18
3	21	100.0	195	2	US-08-485-657A-14
4	21	100.0	195	5	PCR-US95-02303-14
5	17.4	82.9	633	3	US-08-950-925-1
6	17.4	82.9	651	2	US-08-961-858-1
7	17.4	82.9	651	2	US-08-961-858-3
8	17.4	82.9	651	3	US-09-089-593-1
9	17.4	82.9	651	3	US-09-089-593-3
10	17.4	82.9	651	3	US-08-993-380-5
11	17.4	82.9	1173	3	US-08-993-380-3
12	16.8	80.0	330	2	US-08-980-060-9
13	16.8	80.0	330	4	US-09-307-185-9
14	16.8	80.0	334	4	US-09-307-185-90
15	16.8	80.0	1806	2	US-08-980-060-1
16	16.8	80.0	1806	4	US-09-307-185-1
17	16.8	80.0	3459	2	US-08-980-060-3
18	16.8	80.0	3459	4	US-09-307-185-3
19	16.4	78.1	1051	4	US-09-008-697A-21
20	16.2	77.1	497	4	US-09-247-155-41
21	16.2	77.1	550	4	US-08-998-416-18
22	16.2	77.1	688	5	PCR-US93-11310-9
23	16.2	77.1	975	6	5340934-10
24	16.2	77.1	975	6	5340934-9
25	16.2	77.1	1677	2	US-08-684-101-1
26	16.2	77.1	1677	4	US-09-205-814-1
27	16.2	77.1	1677	4	US-09-205-814-1

28	15.8	75.2	361	4	US-09-387-212-5	Sequence 5, Appl 1
29	15.8	75.2	2674	4	US-09-817-180-1	Sequence 1, Appl 1
30	15.8	75.2	15297	4	US-09-817-180-3	Sequence 3, Appl 1
31	15.4	73.3	45	1	US-07-885-689A-7	Sequence 7, Appl 1
32	15.4	73.3	576	1	US-07-885-689A-28	Sequence 28, Appl 1
33	15.4	73.3	576	1	US-08-093-383-4	Sequence 4, Appl 1
34	15.4	73.3	579	1	US-07-885-689A-35	Sequence 35, Appl 1
35	15.4	73.3	579	1	US-07-885-689A-37	Sequence 37, Appl 1
36	15.4	73.3	600	6	5514646-37	Patent No. 5514646
37	15.4	73.3	601	1	US-07-764-655D-6	Sequence 6, Appl 1
38	15.4	73.3	601	1	US-07-801-164A-3	Sequence 3, Appl 1
39	15.4	73.3	603	1	US-07-764-655D-7	Sequence 7, Appl 1
40	15.2	72.4	423	1	US-08-470-179-35	Sequence 35, Appl 1
41	15.2	72.4	423	1	US-08-470-179-51	Sequence 51, Appl 1
42	15.2	72.4	423	1	US-08-470-179-53	Sequence 53, Appl 1
43	15.2	72.4	423	1	US-08-470-179-76	Sequence 76, Appl 1
44	15.2	72.4	423	1	US-08-470-179-80	Sequence 80, Appl 1
45	15.2	72.4	423	1	US-08-470-179-142	Sequence 142, App

#### ALIGNMENTS

RESULT 1  
US-08-485-657A-19/c  
Sequence 19, Application US/08485657A  
Patent No. 5942389  
GENERAL INFORMATION:  
APPLICANT: Kirschling, Deborah J  
APPLICANT: Gudkov, Andrei  
TITLE OF INVENTION: Genes and Genetic Elements Associated  
TITLE OF INVENTION: with sensitivity to Cisplatin  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
STREET: 300 South Wacker Drive, 32nd Floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,657A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5942389nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 93,354-N  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX:  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 181 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-485-657A-19  
Query Match 100.0%, Score 21; DB 2; Length 181;  
Best Local Similarity 100.0%; Pred. No. 0.53;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 cagcagcagagcttcatacat 21

```
Db      73 CAGCAGCAGAGTCTTCATCAT 53
|||||
RESULT 2
PCT-US95-02303-18/C
; Sequence 18, Application PC/TUS9502303
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Genes and Genetic Elements Associated
; TITLE OF INVENTION: With Sensitivity to Cisplatin
; NUMBER OF SEQUENCES: 25
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02303
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US95-02303-18

Query Match      100.0%; Score 21; DB 5; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 21; Conservative 0; Indels 0; Gaps 0;

QY      1 cagcagcagagcttcacatc 21
      |||
Db      73 CAGCAGCAGAGTCTTCATCAT 53

RESULT 3
US-08-485-657A-14/C
; Sequence 14, Application US/08485657A
; Patent No. 5942389
; GENERAL INFORMATION:
; APPLICANT: Kirschling, Deborah J
; APPLICANT: Gudkov, Andrei
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Genes and Genetic Elements Associated
; TITLE OF INVENTION: With Sensitivity to Cisplatin
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.657A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5942389nan, Kevin E
; REGISTRATION NUMBER: 35.303
; REFERENCE/DOCKET NUMBER: 93.354-N
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:

; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-485-657A-14

Query Match      100.0%; Score 21; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 cagcagcagagcttcacatc 21
      |||
Db      87 CAGCAGCAGAGTCTTCATCAT 67

RESULT 5
US-08-950-925-1/C
; Sequence 1, Application US/08950925
; Patent No. 6072040
; GENERAL INFORMATION:
; APPLICANT: Dave, Kirti I.
; APPLICANT: Botvanszki, Janos
; APPLICANT: Sitar, Eva
; TITLE OF INVENTION: Stabilized Conjugates of Uncomplexed
; TITLE OF INVENTION: Subunits of Multimeric Proteins
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

```

:
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/950,925
: FILING DATE:
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Axford, Laurie A.
: REGISTRATION NUMBER: 35,053
: REFERENCE/DOCKET NUMBER: 32260-20004.20
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-813-5600
: TELEFAX: 650-494-0792
: TELEX: 706141
: INFORMATION FOR SEQ. ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 633 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: Coding Sequence
: LOCATION: 1..630
: OTHER INFORMATION:
:
: US-08-950-925-1
:
Query Match      82.9%; Score 17.4; DB 3; Length 633;
Best Local Similarity 94.7%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 cagcagcagagcttcac 19
Db 162 CAGCAGCAGAGCTTCAGC 144

RESULT 6
US-08-961-858-1/C
: Sequence 1, Application US/08961858
: Patent No. 5834210
: GENERAL INFORMATION:
: APPLICANT: Liu, Shigui
: APPLICANT: Shi, Qiuwei
: TITLE OF INVENTION: STABLE TROPONIN SUBUNITS AND COMPLEXES
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauder & Jackson
: STREET: 411 Hackensack Avenue, 4th Floor
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,858
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 1112-1-044 CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-487-5800
: TELEFAX: 201-343-1684
: TELEX: 133521
```

```

:
: INFORMATION FOR SEQ. ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 651 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
:
: US-08-961-858-1
```

```

Query Match      82.9%; Score 17.4; DB 2; Length 651;
Best Local Similarity 94.7%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 cagcagcagagcttcac 19
Db 180 CAGCAGCAGAGCTTCAGC 162
```

```

RESULT 7
US-08-961-858-3/C
: Sequence 3, Application US/08961858
: Patent No. 5834210
: GENERAL INFORMATION:
: APPLICANT: Liu, Shigui
: APPLICANT: Shi, Qiuwei
: TITLE OF INVENTION: STABLE TROPONIN SUBUNITS AND COMPLEXES
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Klauder & Jackson
: STREET: 411 Hackensack Avenue, 4th Floor
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,858
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 1112-1-044 CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-487-5800
: TELEFAX: 201-343-1684
: TELEX: 133521
: INFORMATION FOR SEQ. ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 651 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: DESCRIPTION: native form
: HYPOTHETICAL: NO
:
: US-08-961-858-3

Query Match      82.9%; Score 17.4; DB 2; Length 651;
Best Local Similarity 94.7%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 cagcagcagagcttcac 19
Db 180 CAGCAGCAGAGCTTCAGC 162
```

```
RESULT 8
US-09-089-593-1/C
; Sequence 1, Application US/09089593
; Patent No. 6060278
; GENERAL INFORMATION:
; APPLICANT: Liu, Shigu1
; APPLICANT: Shi, Qianwei
; TITLE OF INVENTION: STABLE TROPONIN SUBUNITS AND COMPLEXES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30.
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/089,593
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,858
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1112-1-044 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-09-089-593-1

Query Match      82.9%; Score 17.4; DB 3; Length 651;
Best Local Similarity 94.7%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 cagcagcagagctctcacc 19
      |||
Db      180 CAGCAGCAGAGCTCTCACC 162

RESULT 9
US-09-089-593-3/C
; Sequence 3, Application US/09089593
; Patent No. 6060278
; GENERAL INFORMATION:
; APPLICANT: Liu, Shigu1
; APPLICANT: Shi, Qianwei
; TITLE OF INVENTION: STABLE TROPONIN SUBUNITS AND COMPLEXES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue, 4th Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/089,593
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,858
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1112-1-044 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
DESCRIPTION: native form
HYPOTHETICAL: NO
US-09-089-593-3

Query Match      82.9%; Score 17.4; DB 3; Length 651;
Best Local Similarity 94.7%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 cagcagcagagctctcacc 19
      |||
Db      180 CAGCAGCAGAGCTCTCACC 162

RESULT 10
US-08-993-380-5/C
; Sequence 5, Application US/08993380B
; Patent No. 6077676
; GENERAL INFORMATION:
; APPLICANT: Song, Qian-Li
; APPLICANT: Shi, Qianwei
; TITLE OF INVENTION: SINGLE CHAIN POLYPEPTIDES COMPRISING TROPONIN I AND
; FILE REFERENCE: 1112-1-053
; CURRENT APPLICATION NUMBER: US/08/993,380B
; CURRENT FILING DATE: 1997-12-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Homo sapiens (modified)
; US-08-993-380-5

Query Match      82.9%; Score 17.4; DB 3; Length 651;
Best Local Similarity 94.7%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 cagcagcagagctctcacc 19
      |||
Db      180 CAGCAGCAGAGCTCTCACC 162

RESULT 11
US-08-993-380-3/C
```

; Sequence 3, Application US/08993380B  
; Patent No. 6077676  
; GENERAL INFORMATION:  
; APPLICANT: Shi, Qian-qi  
; TITLE OF INVENTION: SINGLE CHAIN POLYPEPTIDES COMPRISING TROPONIN C  
; FILE REFERENCE: 1112-1-053  
; CURRENT APPLICATION NUMBER: US/08/993,380B  
; CURRENT FILING DATE: 1997-12-18  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1173  
; TYPE: DNA  
; ORGANISM: Homo sapiens (modified)  
US-08-993-380-3

\* Query Match 82.9%; Score 17.4; DB 3; Length 1173;  
Best Local Similarity 94.7%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 cagcagcaggtcttcac 19  
|||||  
DB 162 CAGCAGCAGGCTTCAC 144

RESULT 12  
US-08-980-060-9/c  
; Sequence 9, Application US/08980060  
; Patent No. 5965421  
; GENERAL INFORMATION:  
; APPLICANT: NI, JIAN  
; APPLICANT: FENG, PING  
; APPLICANT: MUZIO, MARTA  
; APPLICANT: DIXIT, VISHVA M.  
; TITLE OF INVENTION: HUMAN IRAK-2  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/980,060  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEFFE, ERIC K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 1488.1010000/EKS/AJK  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 330 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-980-060-9

Query Match 80.0%; Score 16.8; DB 2; Length 330;

Best Local Similarity 90.0%; Pred. No. 42;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 agcagcaggtcttcacat 21  
|||||  
DB 111 AGCAGGTGAGCTTCATCAT 92

RESULT 13  
US-09-307-185-9/c  
; Sequence 9, Application US/09307185  
; Patent No. 6222019  
; GENERAL INFORMATION:  
; APPLICANT: NI, JIAN  
; APPLICANT: FENG, PING  
; APPLICANT: MUZIO, MARTA  
; APPLICANT: DIXIT, VISHVA M.  
; TITLE OF INVENTION: HUMAN IRAK-2  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/307,185  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/980,060  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEFFE, ERIC K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 1488.1010000/EKS/AJK  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 330 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-307-185-9

Query Match 80.0%; Score 16.8; DB 4; Length 330;  
Best Local Similarity 90.0%; Pred. No. 42;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 agcagcaggtcttcacat 21  
|||||  
DB 111 AGCAGGTGAGCTTCATCAT 92

RESULT 14  
US-09-227-357-90/c  
; Sequence 90, Application US/09227357  
; Patent No. 6342581  
; GENERAL INFORMATION:  
; APPLICANT: Fischer et al.  
; TITLE OF INVENTION: 123 Human Secreted Proteins  
; FILE REFERENCE: P2010P1  
; CURRENT APPLICATION NUMBER: US/09/227,357

```

CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 90
LENGTH: 334
TYPE: DNA
ORGANISM: Homo sapiens

```

```

FEATURE:
NAME/KEY: SITE
LOCATION: (321)
OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-90

Query Match      80.0%; Score 16.8; DB 4; Length 334;
Best Local Similarity 90.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2  accagcagagcttcacat 21
      |||||  |||||
Db      95  accagcagagcttcacat 76

RESULT 15
US-08-980-060-1/c
Sequence 1, Application US/08980060
Patent No. 5965421
GENERAL INFORMATION:
APPLICANT: NI, JIAN
APPLICANT: FENG, PING
APPLICANT: MUZIO, MARTA
APPLICANT: DIXIT, VISHVA M.
TITLE OF INVENTION: HUMAN IRAK-2
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,060
FILING DATE: Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEFFE, ERIC K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.1010000/EKS/AJK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1806 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 34..1803
US-08-980-060-1

Query Match      80.0%; Score 16.8; DB 2; Length 1806;
Best Local Similarity 90.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2  accagcagagcttcacat 21
      |||||  |||||
Db      1157  accagcagagcttcacat 1138

```

Mon May 20 10:17:54 2002

us-09-913-325-4.rni

Page 7

Search completed: May 17, 2002, 16:06:10  
Job time: 12744 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 14:16:28 ; Search time 447.74 Seconds  
(without alignments)  
80.527 Million cell updates/second

```
Title: US-09-913-325-4
Perfect score: 21
Sequence: 1 cagcagcagagtccttcacat 21
```

Scoring table: IDENTITY\_NUC

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```

post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

1:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
N_Geneseq	323802.*
2:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
23:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
24:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	21	100.0	21	AA94426		Human testosterone
2	21	100.0	195	AA00416		Genetic suppressor
3	21	100.0	275	AA006940		Human secreted protein
4	21	100.0	366	AA944824		Human ovarian cancer
5	21	100.0	491	AA003751		Human secreted protein
6	21	100.0	512	AA889150		EST clone CG1
7	21	100.0	572	AA442136		Human normal bladder
8	21	100.0	922	AA543857		Human secreted protein
9	21	100.0	1568	AA544948		cDNA encoding novel

c	10	21	100.0	1651	12	AA011503	Cytolysis inhibitor
c	11	21	100.0	1678	22	AAH23086	Osteoarthritis fls
c	12	17.8	84.8	50	20	AAK24790	Interleukin-18 bin
c	13	17.8	84.8	417	17	AAT64832	Coral cobra venom
c	14	17.8	84.8	418	17	AAT64834	Coral cobra venom
c	15	17.8	84.8	426	17	AAT64833	Coral cobra venom
c	16	17.8	84.8	446	21	AAAC4717	Arabidopsis thailia
c	17	17.8	84.8	448	21	AAAC10235	Human secreted pro
c	18	17.8	84.8	460	17	AAT64831	Coral cobra venom
c	19	17.8	84.8	552	23	ABL10237	Drosophila melanog
c	20	17.8	84.8	960	21	AAAC45936	Arabidopsis thailia
c	21	17.8	84.8	963	21	AAAC45693	Arabidopsis thailia
c	22	17.8	84.8	1201	21	AAAC45344	Arabidopsis thailia
c	23	17.8	84.8	1204	22	AAAC45136	Arabidopsis thailia
c	24	17.8	84.8	3050	23	ABL10236	cDNA encoding nove
c	25	17.4	82.9	236	22	AAH51710	Drosophila melanog
c	26	17.4	82.9	300	21	AAAS3659	Human heart specif
c	27	17.4	82.9	480	23	AAAS87303	Human tropomlin I c
c	28	17.4	82.9	633	21	AAA65127	DNA encoding novel
c	29	17.4	82.9	651	19	AAV06230	Human tropomlin I c
c	30	17.4	82.9	651	20	AAV67262	DNA encoding modif
c	31	17.4	82.9	651	20	AAV67264	Modified human car
c	32	17.4	82.9	651	21	AAAS39301	Native human card
c	33	17.4	82.9	651	21	AAAS39302	Recombinant modifi
c	34	17.4	82.9	669	19	AAV04221	Human cardiac trop
c	35	17.4	82.9	953	19	AAAS87304	DNA encoding modif
c	36	17.4	82.9	960	19	AAV04228	DNA encoding novel
c	37	17.4	82.9	1119	19	AAV04225	DNA encoding card
c	38	17.4	82.9	1173	20	AAVX73382	DNA encoding card
c	39	17.4	82.9	1694	22	AAH57378	Human cardiac trop
c	40	17.4	82.9	1694	22	AAH57378	Human heart cell s
c	41	17.2	81.9	181	21	AAAC24475	Human heart cell s
c	42	16.8	80.0	334	20	AAAT27390	Human secreted pro
c	43	16.8	80.0	478	22	AAI28118	Human breast cance
c	44	16.8	80.0	691	21	AAAC36959	Arabidopsis thailia
c	45	16.8	80.0	719	21	AAAC54096	Arabidopsis thailia

## ALIGNMENTS

RESULT	1
AAAA94226	
ID	AAAA94226 standard; DNA: 21 BP.
XX	
AC	AAAA94226;
DT	12-JAN-2001 (first entry)
XX	
DE	Human testosterone-repressed prostate message-2 antisense oligo #2.
XX	
XX	Human; testosterone-repressed prostate message-2; TRPM-2; cluster1n;
KW	sulfated glycoprotein-2; SGP-2; cancer; antisense oligonucleotide; ss.
XX	
OS	Homo sapiens.
XX	
PN	W020004937-A2.
XX	
PD	31-AUG-2000.
XX	
PF	25-FEB-2000; 2000WO-US04875.
XX	
PR	26-FEB-1999; 99US-0121726.
XX	
PA	(UYBR-) UNIV BRITISH COLUMBIA.
XX	
PV	Gleave M, Rennie PS, Miyake H, Nelson C;
XX	
PI	WPI; 2000-533132/48.
DR	
PT	Treating prostatic tumors and renal cancers by antisense inhibition of
XX	
PT	the testosterone-repressed prostate messenger-2 gene .
XX	

PS Claim 3; Page 36; 38pp; English.

XX The present sequence is an antisense oligonucleotide directed at the  
 CC human testosterone-repressed prostate message-2 (TRPM-2), also known as  
 CC clusterin, sulfated glycoprotein-2 or SCP-2). The sequence was shown to  
 CC promote the regression of tumours, and oligonucleotides directed  
 CC at human TRPM-2 can be used in the treatment of tumour cells expressing  
 CC the TRPM-2 gene. These include prostate cancer, renal cell cancer and  
 CC some breast cancer cells. In addition to this, they also increase the  
 CC chemosensitivity of the cells, meaning that conventional chemotherapy is  
 CC more effective.

XX Sequence 21 BP; 6 A; 6 C; 4 G; 5 T; 0 other;

Query Match 100.0%; Score 21; DB 21; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cagcagcagagctctcatcat 21  
 ||||||||||||||||  
 Db 1 cagcagcagagctctcatcat 21

RESULT 2

AT00416/c  
 ID AT00416 standard; cDNA; 195 BP.

AC AT00416;

DT 26-MAR-1996 (first entry)

XX Genetic suppressor element HL7.1.

XX Genetic suppressor element; GSE; platinum-based drug; cisplatin;  
 KM chemotherapy; HL7.1; testosterone-repressed prostatic message-2; TRPM-2;  
 KM ss.

XX Synthetic.

XX WO9522612-A2.

XX 24-AUG-1995.

XX 22-FEB-1995; 95MO-US02303.

XX 22-FEB-1994; 94US-0199300.

XX (UNII ) UNIV ILLINOIS FOUND.

XX Gudkov A, Kirschling DJ, Roninson IB;

XX WPI; 1995-302718/39.

PT Genetic suppressor elements which confer resistance to  
 PT platinum-based drugs, eg. cisplatin, on cancer cells - useful for  
 PT enhancement of chemotherapy, and for diagnosis of resistance to  
 PT these drugs.

XX Claim 14; Fig 17; 75pp; English.

XX The sequences represented by AT00405-T00418 are genetic suppressor  
 CC elements (GSEs). This sequence represents GSE HL6.10. This sequence  
 CC shows homology to the cDNA encoding testosterone-repressed prostatic  
 CC message-2 (TRPM-2). These sequences were obtained from a cDNA library  
 CC derived from the total cDNA of a cisplatin sensitive cell. Genetic  
 CC suppressor elements confer resistance to platinum-based drugs (Pds),  
 CC such as cisplatin. These functional GSEs can then be used to create  
 CC probes for the parent gene. The probes can then be used in a method of  
 CC measuring the level of GSE gene expression. The GSEs can be used in  
 CC methods of diagnosis of resistance to Pds by measuring the level of  
 CC expression of GSE genes. The GSEs are also used in methods to overcome  
 CC resistance to Pds in cancer cells. The GSEs (or fragments of them) can

CC be used to inhibit the function of genes associated with sensitivity to  
 CC Pds. For enhancement of chemotherapy, a GSE can be transferred (either  
 CC alone or with another gene) on an expression vector into blood  
 CC progenitor cells from a cancer patient. The cells are returned to the  
 CC patients circulation and allowed to repopulate the blood before  
 CC aggressive chemotherapy is carried out (using higher cisplatin  
 CC concentrations than normal), this will thereby avoid toxic side effects  
 CC to the immune system as the blood cells will be GSE resistant.

XX Sequence 195 BP; 46 A; 45 C; 71 G; 33 T; 0 other;

Query Match 100.0%; Score 21; DB 16; Length 195;  
 Best Local Similarity 100.0%; Pred. No. 3.3;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cagcagcagagctctcatcat 21  
 ||||||||||||||||  
 Db 87 CAGCAGCAGAGCTCTCATCAT 67

RESULT 3

AAC06940/c  
 ID AAC06940 standard; cDNA; 275 BP.

AC AAC06940;

DT 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 11015.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KM gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GSEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 11015; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
 CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors.

XX Sequence 275 BP; 59 A; 88 C; 85 G; 43 T; 0 other;

Query Match 100.0%; Score 21; DB 21; Length 275;  
 Best Local Similarity 100.0%; Pred. No. 3.4;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcatcat 21  
 |||  
 DB 128 CAGCAGCAGAGCTTTCATCAT 108

## RESULT 4

AAf94824/C  
 ID AAF94824 standard; cDNA: 396 BP.

XX AAF94824;

XX 23-MAY-2001 (first entry)

XX Human ovarian cancer associated coding sequence SEQ ID NO: 15.

XX Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.

XX Homo sapiens.

XX WO200118046-A2.

XX 15-MAR-2001.

XX 08-SEP-2000; 2000WO-US24827.

XX 10-SEP-1999; 99US-0394374.

XX 01-MAY-2000; 2000US-0561778.

XX 15-AUG-2000; 2000US-0640173.

XX 07-SEP-2000; 2000US-0656668.

XX (CORI-) CORIXA CORP.

XX Xu J, Stolk JA;

XX WPI; 2001-211395/21.

XX Isolated polypeptides associated with ovarian carcinomas, and the

XX nucleic acids that encode them, useful for the prevention diagnosis and

XX treatment of ovarian cancers -

XX Claim 18; Page 120; 189pp; English.

XX The present invention provides a number of coding sequences and proteins,

XX the over-expression of which is associated with ovarian carcinoma/cancer.

XX These can be used in the diagnosis, treatment and prevention of ovarian

XX cancer, optionally by gene therapy or in the form of a vaccine. The

XX present sequence is an example of one of these sequences.

XX Sequence 396 BP; 129 A; 83 C; 121 G; 62 T; 1 other;

## Query Match

Best Local Similarity 100.0%; Score 21; DB 22; Length 396;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagcagcagagcttcatcat 21  
 |||  
 DB 79 CAGCAGCAGAGCTTTCATCAT 59

## RESULT 5

AAc03751/C  
 ID AAC03751 standard; cDNA: 491 BP.

XX AAC03751;

XX 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 3749.  
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EPI033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST ) GENSET.

XX Dumas Milne Edwards J, Duclet A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG03745.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for

XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for

XX diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 3749; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from

XX cDNAs encoding secreted proteins. An ORF has been identified within the

XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs

XX derived from 30 different tissues. EST sequences usually correspond

XX mainly to the 3' untranslated region (UTR) of the mRNA because they are

XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not

XX well suited for isolating cDNA sequences derived from the 5' ends of

XX mRNAs and even in those cases where longer cDNA sequences have been

XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from

XX cDNAs with intact 5' ends and can therefore be used to obtain full length

XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,

XX gene therapy and chromosome mapping procedures. They are used to obtain

XX upstream regulatory sequences and to design expression and secretion

XX vectors.

XX Sequence 491 BP; 149 A; 114 C; 149 G; 78 T; 1 other;

QY 1 cagcagcagagcttcatcat 21  
 |||  
 DB 128 CAGCAGCAGAGCTTTCATCAT 108

## RESULT 6

AAV89150/C  
 ID AAV89150 standard; cDNA: 512 BP.

XX AAV89150;

XX 15-FEB-1999 (first entry)

XX EST clone CG1.

XX Human; secreted protein; expressed sequence tag; EST; haematopoiesis;

XX tissue growth; activin; inhibitor; chemotaxis; chemokinesis; haemostatic;

XX receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;

XX gene therapy; ss.

XX Homo sapiens.

XX W09845436-A2.

```
XX 15-OCT-1998.
PD
XX
XX 10-APR-1998; 98WO-US06955.
PF
XX
XX 10-APR-1997; 97US-0638821.
PR
XX
XX (GENY ) GENETICS INST INC.
PA
XX
XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Werberg D;
PI Racle LA, Spaulding V, Treacy M;
XX WPI; 1999-070077/06.
XX
XX New polynucleotides encoding human secreted proteins - derived from
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT ovary, pituitary, retina and colon cDNA libraries.
XX
XX Claim 1: Page 126: 618pp; English.
PS
XX
XX The present sequence represents a human expressed sequence tag (EST).
CC The polynucleotide, which is a secreted EST, and the encoded protein
CC are predicted to have useful biological activities which would make
CC them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals, although no supporting data is
CC given. Suggested activities include nutritional activity, immune
CC stimulating or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The polynucleotide may also be useful for gene therapy.
XX
XX Sequence 512 BP; 154 A; 114 C; 150 G; 94 T; 0 other;
SQ

Query Match 100.0%; Score 21; DB 20; Length 512;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cagcagcagagctctcatcat 21
DB 55 CAGCAGCAGAGCTCTCATCAT 35

RESULT 7
AA42136/C
ID AA42136 standard; CDNA; 572 BP.
XX
XX AA42136;
AC
XX
XX 31-JAN-2000 (first entry)
DT
XX
XX Human normal bladder tissue cDNA derived EST 15.
DE
XX
XX Human; bladder; treatment; EST; expressed sequence tag; cytostatic;
KW cancer; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX DE19818620-A1.
PN
XX
XX 28-OCT-1999.
PD
XX
XX 21-APR-1998; 98DE-1018620.
PF
XX
XX 21-APR-1998; 98DE-1018620.
PR
XX
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
PA
XX
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
PI WPI; 1999-602416/52.
XX
```

```
XX
XX New polypeptides and their nucleic acids, useful for treatment of
PT bladder tumour and identification of therapeutic agents
XX
XX Claim 3; Page 158; 366pp; German.
PS
XX
XX This invention describes novel polypeptide fragment sequences (I) and
CC their encoding nucleic acids (II) which are highly expressed in normal
CC bladder tissue and have cytostatic activity. (II) are used for
CC recombinant expression of (I) and to isolate complete genes. (I) are
CC used to identify agents suitable for the treatment of bladder tumours,
CC to directly treat this form of cancer (including expression from gene
CC therapy vectors), or are used in a preparation for cancer treatment. (I)
CC is also used for the generation of specific antibodies. (II) are
CC identified by assembling ESTs (expressed sequence tags) from a
CC particular tissue type before comparison of expression patterns. This
CC allows a significantly longer fragment of the gene to be revealed, and
CC therefore reduces the number of failures because of ESTs from different
CC libraries representing different parts of the same unknown gene
CC distorting the estimated frequency of occurrence in a particular tissue.
CC AA42122-242248 represent EST fragments derived from a human normal
CC bladder tissue cDNA library which encode the protein fragments
CC represented in AAY60329-Y60591.
XX
XX Sequence 572 BP; 161 A; 128 C; 186 G; 97 T; 0 other;
SQ

Query Match 100.0%; Score 21; DB 20; Length 572;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 cagcagcagagctctcatcat 21
DB 196 CAGCAGCAGAGCTCTCATCAT 176

RESULT 8
AAA43857/C
ID AAA43857 standard; CDNA; 922 BP.
XX
XX AAA43857;
AC
XX
XX 21-AUG-2000 (first entry)
DT
XX
XX Human secreted expressed sequence tag SEQ ID NO:432.
DE
XX
XX Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;
KW expressed sequence tag; EST; probe; chemotactic; proliferative;
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KW antiviral; antidiabetic; antiaschmatic; vulnary; antiparkinsonian;
KW anticancer; osteoprotective; neuroprotective; noctropic; antipsoriatic;
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KW central nervous system disorder; Alzheimer's disease; stroke;
KW parkinson's disease; Huntington's disease; coagulation disorder;
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KW tumour; infection; depression; psoriasis; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200021991-A1.
PN
XX
XX 20-APR-2000.
PD
XX
XX 15-OCT-1999; 99WO-US24206.
PF
XX
XX 15-OCT-1999; 99WO-US24206.
PR
XX
XX 15-OCT-1998; 98US-0104436.
PA
XX
XX (GENY ) GENETICS INST INC.
XX
```

PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Bowman MR;  
 DR WPI: 2000-317938/27.  
 XX  
 PT Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (seSTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders -  
 XX  
 PS Claim 1: Page 316; 803pp; English.  
 XX  
 CC AAA43426 to AAA45925 represent specifically claimed secreted expressed  
 CC sequence tags (seSTs), isolated from human, mouse, chicken and rat  
 CC tissue sources. The seSTs can have a range of activities depending on  
 CC the tissues they were isolated from. The activities include:  
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
 CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;  
 CC antiallergic; vulnerrary; antiulcer; osteopathic; neuroprotective;  
 CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
 CC anticonvulsant; and antidepressant. The seSTs can be used for gene  
 CC therapy and in vaccines. The seSTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the seSTs. Proteins encoded by the seSTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
 CC in the exemplification of the present invention.  
 XX  
 SQ Sequence 922 BP; 243 A; 265 C; 244 G; 166 T; 4 other;  
 XX  
 Query Match 100.0%; Score 21; DB 21; Length 922;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 cagcagcagagcttcacatc 21  
 Db 52 CAGCAGCAGAGCTTCATCAT 32  
 XX  
 RESULT 9  
 ID AAS44948 standard; cDNA: 1568 BP.  
 XX  
 AC AAS44948;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE cDNA encoding novel human secretory protein, Seq ID No 29.  
 XX  
 KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;  
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;  
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;  
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;  
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;  
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;  
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;  
 KW fertility; analgesic; pain; antigen; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200166689-A2.  
 XX  
 PD 13-SBP-2001.

PF 05-MAR-2001; 2001WO-US04942.  
 XX  
 XX 07-MAR-2000; 2000US-0519705.  
 PR 19-MAY-2000; 2000US-0574454.  
 PR 17-JUN-2000; 2000US-0596193.  
 PR 14-JUL-2000; 2000US-0616647.  
 PR 19-SEP-2000; 2000US-0665363.  
 PR 20-OCT-2000; 2000US-0693267.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;  
 PI Zhao QX, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;  
 DR WPI: 2001-589934/66.  
 DR P-PSDB; AAU28048.  
 XX  
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries  
 PT prepared from various human tissues, for diagnosis and treatment of  
 PT cancer, neurological, inflammatory, and autoimmune disorders -  
 XX  
 PS Claim 1: SEQ ID No 29; 107pp; English.  
 XX  
 CC The invention relates to novel isolated human secreted polypeptides (I)  
 CC and polynucleotides (II). (I) and (II) are useful for treating  
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,  
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is  
 CC involved in increasing haematopoiesis, stem cell survival, bone growth  
 CC and remodeling. (I), (II) and modulators of (II) are useful for  
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for  
 CC creating transgenic animals useful for studying the in vivo activities of  
 CC the polypeptide as well as for studying modulators of the polypeptides.  
 CC (I) induces the proliferation of neural cells and regeneration of nerve  
 CC and brain tissue and is useful for the treatment of central and  
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,  
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral  
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic  
 CC activity, regulation of haematopoiesis and is useful for treating myeloid  
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
 CC tissue growth, and in tissue repair, healing of burns, incisions,  
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative  
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for  
 CC gut protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues, various immune deficiencies and  
 CC disorders including severe combined immunodeficiency (SCID), bacterial or  
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,  
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic  
 CC reactions and conditions, such as asthma or other respiratory problems.  
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,  
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of  
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides  
 CC analgesic effects or other pain reducing effects, immunoglobulin like  
 CC activity and can act as an antigen in a vaccine composition to raise an  
 CC immune response. AAS44920-AAS45295 represent novel human secreted protein  
 CC coding sequences of the invention.  
 XX  
 SQ Sequence 1568 BP; 386 A; 456 C; 434 G; 292 T; 0 other;  
 XX  
 Query Match 100.0%; Score 21; DB 22; Length 1568;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 cagcagcagagcttcacatc 21  
 Db 90 CAGCAGCAGAGCTTCATCAT 70  
 XX  
 RESULT 10  
 ID AAQ11503/C  
 XX  
 AAQ11503 standard; DNA: 1651 BP.

```

AC AAQ11503;
XX
XX 20-JUN-1991 (first entry)
XX
XX Cytolysis inhibitor gene.
DE Cytolysis inhibitor gene.
XX
XX Cytolysis inhibitor; perforin; immunological effector molecule;
KW Interfility; ss.
XX
XX Homo sapiens.
OS
XX
XX Key location/Qualifiers
XX sig.peptide 199..261
XX /*tag= a
XX mat.peptide 262..1542
XX /*tag= b
XX /product= cytolysis inhibitor
XX
XX DE3933850-A.
XX
XX 18-APR-1991.
XX
XX 06-OCT-1989; 89DE-3933850.
XX
XX 06-OCT-1989; 89DE-3933850.
XX
XX (SCHD ) SCHERING AG.
XX
XX Tschopp J, Jenne D;
XX
XX WPI: 1991-118338/17.
XX
XX DNA sequence coding for cytolysis inhibitor - is strong inhibitor
XX of terminal complement protein, e.g. perforin secreted by killer
XX cells
XX
XX Claim 2; Page 8; 15pp; German.
XX
XX Two probes were prepared based on the known partial amino acid
XX sequences of both chains of the Cytolysis inhibitor and used to
XX screen a liver-specific cDNA library. One clone which hybridised
XX positively to both probes was found to contain a 1.7kb BamHI-KpnI
XX fragment. This was inserted into plasmid pGEM4, to give pGEM4/Z1-1.
XX E.coli transformed with the plasmid are deposited under DSM # 5269.
XX The insert has the sequence given in this file.
XX
XX See also AAQ11501 and AAQ11502.
XX
XX Sequence 1651 BP; 405 A; 481 C; 447 G; 318 T; 0 other;
SQ
Query Match 100.0%; Score 21; DB 12; Length 1651;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 cagcagcagagctctcatcat 21
DB 216 CAGCAGCAGAGCTCTCATCAT 196
RESULT 11
AAH23086/C
ID AAH23086 standard; DNA; 1678 BP.
XX
XX AAH23086;
XX
XX 17-SEP-2001 (first entry)
XX
XX Osteoarthritis tissue-derived nucleic acid sequence #16.
DE Osteoarthritis; infectious disorder; Crohn's disease; sepsis; human;
XX wound healing; osteopathic; anti-arthritis; anti-inflammatory;
XX vulnerable; antibacterial; anti-allergic; ds.
XX

```

```

OS Homo sapiens.
XX
XX WO200153531-A2.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US00016.
XX
XX 18-JAN-2000; 2000US-0176523.
XX
XX (PHAA ) PHARMACIA CORP.
XX
XX Phippard D, Vasanthakumur G, Dotson S, Ma X;
XX
XX WPI: 2001-451914/48.
XX
XX Substantially purified protein, polypeptide or their fragments, used to
XX identify a biologically active compound or composition and treat
XX mammalian osteoarthritis
XX
XX Claim 1; Page 102-103; 144pp; English.
XX
XX Sequences AAH23071-23152 represent nucleic acid sequences derived from
XX osteoarthritis tissues. The sequences are useful as probes and for the
XX diagnosis or prognosis of mammalian osteoarthritis. The polynucleotides
XX and polypeptides of the invention are useful for generating diagnostic
XX reagents, as targets for small molecule drug development, generation of
XX therapeutics, and cloning genes. Specific antibodies are used to generate
XX enzyme linked immunosorbent assays for detection of osteoarthritis. The
XX invented molecules can be used to treat osteoarthritis or to analyse the
XX disease-modifying activity of osteoarthritis drugs. Other disorders
XX treatable using the nucleic acid sequences include atopic, inflammatory
XX and infectious disorders e.g. Crohn's disease and sepsis, and wound
XX healing.
XX
XX Sequence 1678 BP; 424 A; 491 C; 450 G; 313 T; 0 other;
SQ
Query Match 100.0%; Score 21; DB 22; Length 1678;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 cagcagcagagctctcatcat 21
DB 90 CAGCAGCAGAGCTCTCATCAT 70
RESULT 12
AAH24790/C
ID AAH24790 standard; DNA; 50 BP.
XX
XX AAH24790;
XX
XX 21-JUN-1999 (first entry)
XX
XX Interleukin-18 binding protein splice variant IL-18BPA PCR primer.
DE Interleukin-18 binding protein; IL-18BP; IL-18BPA; splice variant;
XX human; autoimmune disease; inflammation; diabetes; pancreatitis;
XX rheumatoid arthritis; graft rejection; sepsis; chronic hepatitis;
XX psoriasis; inflammatory bowel disease; multiple sclerosis;
XX ischemic heart disease; ischemic brain injury; gene therapy; PCR;
XX primer; ss.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX
XX WO9909063-A1.
XX
XX 25-FEB-1999.
XX
XX 13-AUG-1998; 98WO-IL00379.
XX

```

PR 22-JUL-1998; 98IL-0125463.  
 PR 14-AUG-1997; 97IL-0121554.  
 PR 27-AUG-1997; 97IL-0121639.  
 PR 29-SEP-1997; 97IL-0121860.  
 PR 06-NOV-1997; 97IL-0122134.  
 XX  
 PA (YEDA ) YEDA RES & DEV CO LTD.  
 XX  
 PI Dinarello C, Kim SH, Novick D, Rubinstein M;  
 XX WPI, 1999-180975/25.  
 DR  
 XX New interleukin-18 binding protein - useful for treating human  
 PT diseases, including autoimmune disease and inflammation  
 PS  
 XX Example 7; Page 39; 100pp; English.  
 CC The present sequence is a reverse primer that was used with a  
 CC forward primer (see AAX24754) in the PCR amplification of the coding  
 CC region of a cDNA clone (see AAX24749) coding for human interleukin-18  
 CC binding protein splice variant IL-18Bpa (see AAW98004). The PCR  
 CC products were cut with XbaI and cloned into the XbaI site of the  
 CC pEF-BOS expression vector. IL-Bp18a was expressed in transfected  
 CC COS7 cells. The invention provides IL-18BP polypeptides capable of  
 CC binding IL-18 and/or modulating and/or blocking IL-18 activity.  
 CC These can be used to treat conditions requiring the protein  
 CC (claimed), such as autoimmune diseases, type I diabetes, rheumatoid  
 CC arthritis, graft rejections, inflammatory bowel disease, sepsis,  
 CC multiple sclerosis, ischemic heart diseases, ischemic brain injury,  
 CC chronic hepatitis, psoriasis, and chronic/acute pancreatitis.  
 XX  
 SQ Sequence 50 BP; 13 A; 8 C; 13 G; 16 T; 0 other;  
 Query Match 84.8%; Score 17.8; DB 20; Length 50;  
 Best Local Similarity 90.5%; Pred: No. 76;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 cagcagcagagcttcacat 21  
 Db 42 CAGCAGCAGGTCATCATCAT 22  
 RESULT 13  
 AAT64832/c  
 ID AAT64832 standard; cDNA; 417 BP.  
 XX  
 AC AAT64832;  
 XX  
 DT 04-SEP-1997 (first entry)  
 XX  
 DE Coral cobra venom protein cDNA clone VI.  
 XX  
 KW Coral cobra; snake venom; toxin; snakebite; antidote; antiserum;  
 KM immunodominant; ss.  
 XX  
 OS Micrurus corallinus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 15..251  
 FT /tag= a  
 FT /product= putative\_toxin  
 FT 15..77  
 FT /tag= b  
 FT /note= "deduced by comparison to homologous sequences"  
 FT 401..406  
 FT /tag= c  
 FT 417  
 FT /tag= d  
 FT polyA\_site  
 FT 417  
 XX BR9403020-A.  
 XX PD 03-SEP-1996.

XX 30-AUG-1994; 94BR-0003020.  
 XX  
 PR 30-AUG-1994; 94BR-0003020.  
 XX  
 PA (BUTA-) FUNDACAO BUTANTAN.  
 PA (HOPL/) HO P L.  
 XX  
 PI Ho PL;  
 XX  
 DR WPI, 1996-434245/44.  
 DR P-PSDB; AAW18368.  
 XX  
 PT Coral cobra venom toxin cDNA - and prepn of the venom by genetic  
 PT engineering or chemical synthesis for prodn. of antiserum to treat  
 PT snake bites  
 PS  
 XX Claim 1; Fig 7(A); 16pp; Portuguese.  
 CC The abundant 07 cDNAs which code for immunodominant toxins from  
 CC the poison glands of the coral cobra (Micrurus corallinus) have been  
 CC identified and cloned. The present sequence is that of cDNA clone  
 CC VI. The sequences from clones VI, V3 and V7 (see AAT64832, AAT64833 and  
 CC AAT64834) are all very similar, with V3 and V7 giving the same  
 CC deduced amino acid sequence. The toxins can be used for preparing  
 CC antisera to treat coral cobra bites.  
 CC (N.B. The numbering shown on figure 7 gives the sequence as being  
 CC 447 nucleotides long).  
 XX  
 SQ Sequence 417 BP; 103 A; 94 C; 91 G; 129 T; 0 other;  
 Query Match 84.8%; Score 17.8; DB 17; Length 417;  
 Best Local Similarity 90.5%; Pred: No. 1e+02; 2; Indels 0; Gaps 0;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 cagcagcagagcttcacat 21  
 Db 32 CAGCAGCAGAGTTCATCTT 12  
 RESULT 14  
 AAT64834/c  
 ID AAT64834 standard; cDNA; 418 BP.  
 XX  
 AC AAT64834;  
 XX  
 DT 04-SEP-1997 (first entry)  
 XX  
 DE Coral cobra venom protein cDNA clone V7.  
 XX  
 KW Coral cobra; snake venom; toxin; snakebite; antidote; antiserum;  
 KM immunodominant; ss.  
 XX  
 OS Micrurus corallinus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 16..252  
 FT /tag= a  
 FT /product= putative\_toxin  
 FT 16..78  
 FT /tag= b  
 FT /note= "deduced by comparison to homologous sequences"  
 FT 402..407  
 FT /tag= c  
 FT 418  
 FT /tag= d  
 FT polyA\_site  
 FT 418  
 XX BR9403020-A.  
 XX PD 03-SEP-1996.  
 XX 30-AUG-1994; 94BR-0003020.

RESULT	15
AA64833/c	
ID	AA64833 standard; cDNA; 426 BP.
XX	
XX	AA64833;
XX	
XX	04-SEP-1997 (first entry)
XX	
DE	Coral cobra venom protein cDNA clone V3.
XX	
XX	Coral cobra; snake venom; toxin; snakebite; antidote; antiserum;
KW	immunodominant; ss.
XX	
OS	Micrurus corallinus.
XX	
XX	
FH	key
FT	Location/Qualifiers
FT	24..260
FT	/*tag= a
FT	/product= putative_toxin
FT	24..86
FT	/*tag= b
FT	/note= "deduced by comparison to homologous sequences
FT	polyA_signal
FT	410..415
FT	/*tag= c
FT	426
FT	polyA_site
FT	/*tag= d
XX	
XX	
XX	BR9403020-A.
XX	
XX	
XX	03-SEP-1996.
XX	
XX	30-AUG-1994;
XX	
XX	94BR-0003020.
XX	
XX	30-AUG-1994;
XX	
XX	94BR-0003020.

Search completed: May 17, 2002, 16:13:54  
Job time: 7046 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 11:35:40 ; Search time 1959.41 Seconds  
(without alignments)  
224.280 Million cell updates/sec

Title: US-09-913-325-4

Perfect score: 21  
Sequence: 1 cagcagcagagcttcacatc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank! :  
1: gb\_da : \*  
2: gb\_hg : \*  
3: gb\_in : \*  
4: gb\_on : \*  
5: gb\_ov : \*  
6: gb\_pat : \*  
7: gb\_ph : \*  
8: gb\_pl : \*  
9: gb\_pr : \*  
10: gb\_ro : \*  
11: gb\_sts : \*  
12: gb\_sy : \*  
13: gb\_un : \*  
14: gb\_vi : \*  
15: em\_da : \*  
16: em\_fun : \*  
17: em\_hum : \*  
18: em\_in : \*  
19: em\_mu : \*  
20: em\_om : \*  
21: em\_or : \*  
22: em\_ov : \*  
23: em\_pat : \*  
24: em\_ph : \*  
25: em\_pl : \*  
26: em\_ro : \*  
27: em\_sts : \*  
28: em\_un : \*  
29: em\_vi : \*  
30: em\_htg\_hum : \*  
31: em\_htg\_inv : \*  
32: em\_htg\_other : \*  
33: em\_htgo\_inv : \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

C	1	21	100.0	396	6	AX093197	AX093197 Sequence
C	2	21	100.0	1648	9	HUMRPM2A	M64722 Human TRPM
C	3	21	100.0	1651	6	A21577	A21577 Blood Plasm
C	4	21	100.0	1651	9	HUMCLI	M25915 Human compl
C	5	21	100.0	1658	9	BC019588	BC019588 Homo sapi
C	6	21	100.0	1676	9	HSCSP40	X147223 Human SP-40
C	7	21	100.0	1678	6	AX202086	AX202086 Sequence
C	8	21	100.0	1684	9	BC010514	BC010514 Homo sapi
C	9	21	100.0	7610	9	HUMRPM2A1	M63376 Human TRPM
C	10	21	100.0	175382	9	AF235104	AF235104 Homo sapi
C	11	21	100.0	177453	2	AC025192	AC025192 Homo sapi
C	12	21	100.0	234431	9	AF311103	AF311103 Homo sapi
C	13	19.4	92.4	1636	4	DOGCP80	M55251 Dog glycopr
C	14	19.4	92.4	1678	4	AF118852	AF118852 Oryctolag
C	15	18.4	87.6	3287	4	AF063235	AF063235 Pneumocys
C	16	18	85.7	1632	4	PICCCIMRA	M84639 Pig complem
C	17	17.8	84.8	264	5	AF023271	AF023271 NaJa sput
C	18	17.8	84.8	264	5	AF023272	AF023272 NaJa sput
C	19	17.8	84.8	278	5	NS086588	U86588 NaJa sputat
C	20	17.8	84.8	278	5	NS086589	U86589 NaJa sputat
C	21	17.8	84.8	278	5	NS086591	U86591 NaJa sputat
C	22	17.8	84.8	278	5	NS086592	U86592 NaJa sputat
C	23	17.8	84.8	278	5	NS086593	U86593 NaJa sputat
C	24	17.8	84.8	278	5	NS086594	U86594 NaJa sputat
C	25	17.8	84.8	278	5	NS086595	U86595 NaJa sputat
C	26	17.8	84.8	278	5	NS086596	U86596 NaJa sputat
C	27	17.8	84.8	278	5	NS086597	U86597 NaJa sputat
C	28	17.8	84.8	283	11	G09801	G09801 human STS C
C	29	17.8	84.8	290	5	AF276223	AF276223 NaJa sput
C	30	17.8	84.8	294	5	AF098923	AF098923 NaJa sput
C	31	17.8	84.8	294	5	AF098924	AF098924 NaJa sput
C	32	17.8	84.8	349	5	ALTOXD	X13372 Alpyrusus 1
C	33	17.8	84.8	365	5	AF142323	AF142323 Bungarus
C	34	17.8	84.8	392	5	DADACSCI	X53410 D.angustice
C	35	17.8	84.8	392	5	DAPLP	X52260 Dendroaspis
C	36	17.8	84.8	408	5	AF276222	AF276222 NaJa sput
C	37	17.8	84.8	418	5	AF197563	AF197563 Micrurus
C	38	17.8	84.8	420	5	AF197565	AF197565 Micrurus
C	39	17.8	84.8	428	5	AF197564	AF197564 Micrurus
C	40	17.8	84.8	452	5	BMKNCB1	X51412 Bungarus mu
C	41	17.8	84.8	467	5	LSEARABU	X16950 Laticauda s
C	42	17.8	84.8	477	5	AB015513	AB015513 Laticauda
C	43	17.8	84.8	480	5	AF026893	AF026893 NaJa sput
C	44	17.8	84.8	483	5	LSEARAR	X02533 Laticauda s
C	45	17.8	84.8	486	5	MC0344067	AJ344067 Micrurus

## ALIGNMENTS

RESULT 1  
LOCUS AX093197/c 396 bp DNA linear PAT 30-MAR-2001  
DEFINITION Sequence 15 from Patent WO0118046.  
ACCESSION AX093197  
VERSION AX093197.1 GI:13509646  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 396)  
AUTHORS Xu, J. and Stol, J.A.  
TITLE Ovarian tumor sequences and methods of use therefor  
JOURNAL Patent: WO 0118046-A 15 15-MAR-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source location/Qualifiers  
1..396  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
misc\_feature 1..396  
BSEQ COUNT 129 a 83 c 121 g 62 t 1 others

## ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cagcagcagagcttcacatc 21  
 ||||||||||||||||  
 Db 79 CAGCAGCAGAGCTTCATCAT 59

## RESULT 2

HUMTRPM2A/c HUMTRPM2A 1648 bp mRNA linear PRI 23-AUG-1996  
 LOCUS Human TRPM-2 mRNA, complete cds.  
 DEFINITION M64722  
 ACCESSION M64722.1 GI:339972  
 VERSION TRPM-2 protein.  
 KEYWORDS Human cDNA to mRNA.  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (sites)  
 Wong, P., Pineault, J., Lakin, J., Taillefer, D., Leger, J., Wang, C.  
 and Tenniswood, M.  
 Genomic organization and expression of the rat TRPM-2 (clusterin)  
 gene, a gene implicated in apoptosis  
 J. Biol. Chem. 268 (7), 5021-5031 (1993)

JOURNAL 93186813  
 MEDLINE 2 (bases 1 to 1648)  
 REFERENCE Wong, P., Taillefer, D., Lakin, J., Pineault, J., Chader, G. and  
 Tenniswood, M.  
 Molecular characterization of human TRPM-2/clusterin, a gene  
 associated with sperm maturation, apoptosis and neurodegeneration  
 Eur. J. Biochem. 221 (3), 917-925 (1994)

TITLE  
 JOURNAL  
 MEDLINE  
 FEATURES 94237156 Location/Qualifiers  
 source 1..1648  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 1..1648  
 /gene="TRPM-2"  
 53..1402  
 /gene="TRPM-2"  
 /codon\_start=1  
 /protein\_id="AAB06508.1"  
 /db\_xref="GI:339973"  
 /translation="MMKTLLEFVGLLTWESGOVLGDOTVSDNLEQNSNGSKYVK  
 EIQAVNGVKQIKLTLEKTEERKTLISNLEAKKKEDALNTRSEETKLELPGVC  
 NETMALMECKPCLOKOTCMKFYARVCRSGSLGVLQEEFLNOSPFPMMNGDRID  
 SLLENDROQTHMDVMODHFRASSIIDLFDQDPTREPOTYHYLPSPHPHREPR  
 PFPSRYRSLMPSPEYEPHFMQOPLHMHQOMDHFHSPAFQCHPEPRF  
 EGGDRTVCREIRHNSGTCLRMKQCDKCRREILSDCSTNNPQAKLRDESLQVA  
 ERLTRKYNELKSTQWKMNTSSLLEQLWEDFNWVSRLANLTQGEDOYLLVTVTASH  
 TSDSDVSGVTEVYVKLFDSDPITVTVPVEVSRKPKFMETVAEKALQEVKKHRE"  
 1627..1632  
 /gene="TRPM-2"  
 /base\_count 408 a 488 c 440 g 312 t

polyA\_signal  
 BASE COUNT 408 a 488 c 440 g 312 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 1648;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cagcagcagagcttcacatc 21  
 ||||||||||||||||  
 Db 73 CAGCAGCAGAGCTTCATCAT 53

## RESULT 3

A21577/c

LOCUS A21577 1651 bp DNA linear PAT 26-JUL-1994  
 DEFINITION blood plasma component having a biological activity of inhibiting  
 cytolysis mediated by a cytolytic protein.

ACCESSION A21577  
 VERSION A21577.1 GI:579601  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1651)

AUTHORS  
 TITLE CYTOLYSIS INHIBITOR PROTEINS (CLI) AND DNA SEQUENCES CODING FOR  
 SAID PROTEINS  
 JOURNAL Patent: WO 9105043-A 3 18-Apr-1991;  
 FEATURES Location/Qualifiers  
 source 1..1651  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 199..1545  
 /codon\_start=1  
 /product="blood plasma component having a biological  
 activity of inhibiting cytolysis mediated by a cytolytic  
 protein"  
 /protein\_id="CA01560.1"  
 /db\_xref="GI:579602"  
 /db\_xref="SWISS-PROT:P10909"  
 /translation="MKTLLEFVGLLTWESGOVLGDOTVSDNLEQNSNGSKYVK  
 IONAVNGVKQIKLTLEKTEERKTLISNLEAKKKEDALNTRSEETKLELPGVCN  
 ETMALMECKPCLOKOTCMKFYARVCRSGSLGVLQEEFLNOSPFPMMNGDRIDS  
 LLENDROQTHMDVMODHFRASSIIDLFDQDPTREPOTYHYLPSPHPHREPR  
 PFPSRYRSLMPSPEYEPHFMQOPLHMHQOAMDHFHSPAFQCHPEPRF  
 GDDRTVCREIRHNSGTCLRMKQCDKCRREILSDCSTNNPQAKLRDESLQVAE  
 RLTTRYNELKSTQWKMNTSSLLEQLWEDFNWVSRLANLTQGEDOYLLVTVTASH  
 TSDSDVSGVTEVYVKLFDSDPITVTVPVEVSRKPKFMETVAEKALQEVKKHRE"

## CDS

BASE COUNT 405 a 481 c 447 g 318 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 1651;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cagcagcagagcttcacatc 21  
 ||||||||||||||||  
 Db 216 CAGCAGCAGAGCTTCATCAT 196

## RESULT 4

HUMCLI/c HUMCLI 1651 bp mRNA linear PRI 27-APR-1993  
 LOCUS Human complement cytolysis inhibitor (CLI) mRNA, complete cds.  
 DEFINITION M25915  
 ACCESSION M25915.1 GI:180619  
 VERSION M25915.1  
 KEYWORDS blood plasma glycoprotein; complement cytolysis inhibitor;  
 secretory glycoprotein; seminal plasma protein; soluble terminal  
 complement complex (C5b-9s) component.  
 SOURCE Human (adult) liver, cDNA to mRNA, clone CLI-1.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1651)  
 Jemne, D.E. and Tschopp, J.  
 Molecular structure and functional characterization of a human  
 complement cytolysis inhibitor found in blood and seminal plasma:  
 identity to sulfated glycoprotein 2, a constituent of rat testis  
 fluid  
 Proc. Natl. Acad. Sci. U.S.A. 86, 7123-7127 (1989)

AUTHORS  
 TITLE  
 JOURNAL  
 MEDLINE  
 COMMENT Draft entry and computer-readable sequence for [1] kindly submitted  
 by D.E.Jemne, 05-JUL-1989.  
 FEATURES Location/Qualifiers  
 source 1..1651

CDs  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
199..1345  
/note="complement cytolysis inhibitor precursor"  
/codon\_start=1  
/db\_xref="GI:180620"  
/translation="MKTLLEPVGLLTVESGOVLDQTVSDNLEMSNOGSKYVKE  
IONAVNGVKQIKTLERKTEKRLTSLNEAKKKEDALNTERSETRKLELPGVC  
ETMALMEBCKPCLKOTCKRYARCRSGSLVGRQLEELNOSPFYEMNGDRIDS  
LLENDROQTHMLDVMODHFRASSTIDELFQDRFTREPQDTHTLPLPRPHFE  
FPKSRIVSLMPSPYEPFLNFMQPELMEIHEAQAMDIFHSPAFQHPPELRE  
GDDRTVCREIRHNSVGLRMKDDCKREILSYDCSTNNPSQAKLRLEDESLQVAE  
RLTRKYNELKSYOMKMLNTSLLEQNEQFNWVSLANLTQGEQYLLRTVTYASH  
SDSDVPSGVTEVYVRLFDSDPTITVTPVEVSRRNPKFMETVAEKALOEYRKKHRE"  
199..261

sig\_peptide  
/note="complement cytolysis inhibitor signal peptide"  
262..876  
mat\_peptide  
/product="complement cytolysis inhibitor a-chain"  
877..1342  
/product="complement cytolysis inhibitor b-chain"  
BASE COUNT 405 a 481 c 447 g 318 t  
ORIGIN Unreported.

Query Match 100.0%; Score 21; DB 9; Length 1651;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 cagcagcagagctctcatcat 21  
|||||  
Db 216 CAGCAGCAGAGCTCTCATCAT 196

RESULT 5  
LOCUS BC019588/c 1658 bp mRNA linear PRI 22-JAN-2002  
DEFINITION Homo sapiens, clusterin (complement lysis inhibitor, SP-40, 40,  
sulfated glycoprotein 2, testosterone-repressed prostate message 2,  
apolipoprotein J), clone MGC:24903 IMAGE:4915444, mRNA, complete  
cds.

ACCESSION BC019588  
VERSION BC019588.1 GI:18043614  
KEYWORDS MGC.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 1658)  
Strausberg, R.  
Direct Submission  
Submitted (19-DEC-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: David N. Louis, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarane, P.H., Garcia, A.M., Lu, X., Huliy, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>  
Series: IRAC Plate: 29 Row: m Column: 21

FEATURES  
source  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4502904.  
Location/Qualifiers  
1..1658

/organism="Homo sapiens"  
/db\_xref="LocusID:1191"  
/db\_xref="taxon:9606"  
/clone="MGC:24903 IMAGE:4915444"  
/tissue\_type="Brain, anaplastic oligodendroglioma with  
1p/19q loss"  
/clone\_1lib="NCI-CGAP\_Brn67"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
49..1398

CDs  
/codon\_start=1  
/product="clusterin (complement lysis inhibitor, SP-40, 40,  
sulfated glycoprotein 2, testosterone-repressed prostate  
message 2, apolipoprotein J)"  
/protein\_id="AAH19588.1"  
/db\_xref="GI:18043615"  
/translation="MKTLLEPVGLLTVESGOVLDQTVSDNLEMSNOGSKYVKN  
EIONAVNGVKQIKTLERKTEKRLTSLNEAKKKEDALNTERSETRKLELPGVC  
ETMALMEBCKPCLKOTCKRYARCRSGSLVGRQLEELNOSPFYEMNGDRIDS  
LLENDROQTHMLDVMODHFRASSTIDELFQDRFTREPQDTHTLPLPRPHFE  
FPKSRIVSLMPSPYEPFLNFMQPELMEIHEAQAMDIFHSPAFQHPPELRE  
EGDDRTVCREIRHNSVGLRMKDDCKREILSYDCSTNNPSQAKLRLEDESLQVAE  
RLTRKYNELKSYOMKMLNTSLLEQNEQFNWVSLANLTQGEQYLLRTVTYASH  
SDSDVPSGVTEVYVRLFDSDPTITVTPVEVSRRNPKFMETVAEKALOEYRKKHRE"  
BASE COUNT 423 a 486 c 439 g 310 t  
ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 1658;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 cagcagcagagctctcatcat 21  
|||||  
Db 69 CAGCAGCAGAGCTCTCATCAT 49

RESULT 6  
LOCUS HSCSP40/c 1676 bp mRNA linear PRI 22-MAR-1995  
DEFINITION Human SP-40, 40 mRNA for complement-associated protein SP-40, 40  
alpha-1 and beta-1 chain.  
X14723  
X14723.1 GI:30250  
KEYWORDS complement-associated protein; serum protein; SP-40, 40 gene.  
human.  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 1676)  
Kirsbaum, L.  
Direct Submission  
Submitted (17-MAR-1989) Kirsbaum L., The University of Melbourne,  
The Preclinical Centre, School of Veterinary Science, Parkville  
Victoria 3052, Australia  
2 (bases 1 to 1676)  
Kirsbaum, L., Sharpe, J.A., Murphy, B., d'Aplice, A.J., Claason, B.,  
Hudson, P., and Walker, I.D.  
Molecular cloning and characterization of the novel, human  
complement-associated protein, SP-40, 40: a link between the  
complement and reproductive systems  
EMBO J. 8 (3), 711-718 (1989)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT  
The sequence overlaps with that reported by Murphy et. al. in J.  
Clin. Invest. 81:1858-1864(1988).  
location/Qualifiers  
1..1676  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

FEATURES  
source

```

/c1one="LK (107)"
/c1one_type="Liver"
/c1one_lib="lambda g111"
48. .113
48. .1397
/note="SP-40,40 prepropeptide (AA -22 to 427)"
/codon_start=1
/protein_id="CA32847.1"
/db_xref="GI:30251"
/db_xref="SWISS-PROT:P10909"
/translation="MKTKLLFVGLLFWESGQVLGDQTVSDNELQMSNOGSKYVVK
EIONAVGVQIKITLTKTEKNEKTLNLEAKKKKEDALNTRSEETLKLEPGVC
NETMALMECKPCLOKOTCMFYARVCSGSLVQRLKEELFNOSPFYMMGDRID
SLENDROOTHMIDVMOHFRASSIIDLEODREFTPQDQTHYLPFSLPHRRPFR
FPKRSRIVRELMPEFSPYEDLNHAFHOPELHIEAOQANDIHFHSPAFQHPFERF
EGDDRTVCREIRHNSTGCLRMKDDCKCREILSYDCSTNNPSQAKLRRLDESIOVA
ERLURKYNELKSYQMKMLNTSSLEQLEDFNVSRLANITQGEQDYLRVTYVASH
TSDSDVPSGYTEVYVKLFPSDPITVTVPVEVSRKPKFMETVAEKALQERYKKHRE"
114. .>114
/note="beta-chain"
114. .1394
/note="SP-40,40 propeptide (AA 1-427)"
729. .1394
/product="mature alpha-chain (AA 205-427)"
1622. .1627
/note="poc. polyA signal"
BASE COUNT      436 a      488 c      437 g      315 t
ORIGIN

Query Match      100.0%; Score 21; DB 9; Length 1676;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 cagcagcagagcttcacat 21
|||||
Db      68 CAGCAGCAGAGCTTCATCAT 48

RESULT 7
AX202086/c      1678 bp      DNA      Linear      PAT 30-AUG-2001
LOCUS      AX202086
DEFINITION      Sequence 16 from Patent WO0153531.
ACCESSION      AX202086
VERSION      AX202086.1 GI:15391872
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 1678)
AUTHORS      Philippard,D., Vasanthakumari,G., Dotson,S. and Ma,X.J.
TITLE      Osteoarthritis tissue derived nucleic acids, polypeptides, vectors,
and cells
JOURNAL      Patent: WO 0153531-A 16 26-JUL-2001;
Pharmacia Corporation (US)
FEATURES
source      1. 1678
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      424 a      491 c      450 g      313 t
ORIGIN

Query Match      100.0%; Score 21; DB 6; Length 1678;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 cagcagcagagcttcacat 21
|||||
Db      90 CAGCAGCAGAGCTTCATCAT 70

```

```

RESULT 8
BC010514/c      1684 bp      mRNA      Linear      PRI 12-JUL-2001
LOCUS      BC010514
DEFINITION      Homo sapiens, clusterin (complement lysis inhibitor, SP-40,40,
sulfated glycoprotein 2, testosterone-repressed prostate message 2,
apolipoprotein J), clone MGC:18080 IMAGE:4150452, mRNA, complete
cds
ACCESSION      BC010514
VERSION      BC010514.1 GI:14714740
KEYWORDS      MGC.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 1684)
AUTHORS      Strausberg,R.
TITLE      Direct Submission
JOURNAL      Submitted (10-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
Email: gcgabs-yemail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc..
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalona@bcm.tmc.edu.
Villalona, D.K., Luna, R.A., Haje, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Hollway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNI at: http://image.llnl.gov
Series: IRAC Plate: 19 Row: 0 Column: 12.
Location/Qualifiers
1. 1684
/organism="Homo sapiens"
/db_xref="locusid:1191"
/db_xref="taxon:9606"
/c1one="MGC:18080 IMAGE:4150452"
/tissue_type="Brain, glioblastoma with EGFR amplification"
/c1one_lib="NCI-CGAP_Brn64"
/lab_host="DH10B"
/note="vector: pcMV-SPORT6"
69. .1418
/codon_start=1
/product="clusterin (complement lysis inhibitor, SP-40,40,
sulfated glycoprotein 2, testosterone-repressed prostate
message 2, apolipoprotein J)"
/protein_id="AAH0514.1"
/db_xref="GI:14714741"
/translation="MKTKLLFVGLLFWESGQVLGDQTVSDNELQMSNOGSKYVVK
EIONAVGVQIKITLTKTEKNEKTLNLEAKKKKEDALNTRSEETLKLEPGVC
NETMALMECKPCLOKOTCMFYARVCSGSLVQRLKEELFNOSPFYMMGDRID
SLENDROOTHMIDVMOHFRASSIIDLEODREFTPQDQTHYLPFSLPHRRPFR
FPKRSRIVRELMPEFSPYEDLNHAFHOPELHIEAOQANDIHFHSPAFQHPFERF
EGDDRTVCREIRHNSTGCLRMKDDCKCREILSYDCSTNNPSQAKLRRLDESIOVA
ERLURKYNELKSYQMKMLNTSSLEQLEDFNVSRLANITQGEQDYLRVTYVASH
TSDSDVPSGYTEVYVKLFPSDPITVTVPVEVSRKPKFMETVAEKALQERYKKHRE"
BASE COUNT      433 a      493 c      448 g      310 t
ORIGIN

Query Match      100.0%; Score 21; DB 9; Length 1684;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 cagcagcagagcttcacat 21

```



ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 187453)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 JOURNAL Homo sapiens chromosome 8, clone RP11-509E2  
 TITLE Unpublished  
 COMMENT 2 (bases 1 to 187453)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
 Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,  
 Campiano, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S.,  
 Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
 Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Grant, G., Hages, B., Heatford, A., Horton, L.,  
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
 Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehocsky, J.,  
 Levine, R., Liu, G., Locke, K., MacDonald, P., Marquis, N.,  
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,  
 Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
 Plisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
 Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Teafaye, S., Theodore, J., Tirrell, A., Travers, M., Triggilio, J.,  
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (07-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jan 19, 2001 this sequence version replaced gi:6901239.  
 All repeats were identified using RepeatMasker:  
 Smt, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 -----  
 Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: MIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 -----  
 Project Information  
 Center project name: L5592  
 Center clone name: 509\_E-2  
 -----  
 Summary Statistics  
 Sequencing vector: M13; M7815; 48% of reads  
 Sequencing vector: plasmid; n/a; 52% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 185120 bases at least Q40  
 Consensus quality: 186074 bases at least Q30  
 Consensus quality: 186533 bases at least Q20  
 Insert size: 175000; agarose-fp  
 Insert size: 186853; sum-of-contigs  
 Quality coverage: 8.9 in Q20 bases; agarose-fp  
 Quality coverage: 8.3 in Q20 ba.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 7 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 17773: contig of 17773 bp in length  
 \* 17774 17873: gap of 100 bp  
 \* 17874 20278: contig of 2405 bp in length  
 \* 20279 20378: gap of 100 bp  
 \* 20379 35742: contig of 15364 bp in length  
 \* 35743 35842: gap of 100 bp  
 \* 35843 65413: contig of 29571 bp in length  
 \* 65414 65513: gap of 100 bp  
 \* 65514 112793: contig of 47280 bp in length

FEATURES  
 SOURCE  
 \* 112794 112893: gap of 100 bp  
 \* 112894 150031: contig of 37138 bp in length  
 \* 150032 150131: gap of 100 bp  
 \* 150132 187453: contig of 37322 bp in length.  
 Location/Qualifiers  
 1. 187453  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8"  
 /clone="RP11-509E2"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 1. 17773  
 /note="assembly\_fragment"  
 /vector\_side="left"  
 /vector\_end="SP6"  
 17874. .20278  
 /note="assembly\_fragment"  
 /note="assembly\_fragment"  
 20379. .35742  
 /note="assembly\_fragment"  
 /note="assembly\_fragment"  
 35843. .65413  
 /note="assembly\_fragment"  
 /note="assembly\_fragment"  
 65514. .112793  
 /note="assembly\_fragment"  
 /note="assembly\_fragment"  
 112894. .150031  
 /note="assembly\_fragment"  
 /note="assembly\_fragment"  
 150132. 187453  
 /note="assembly\_fragment"  
 /vector\_side="right"  
 /vector\_end="T7"  
 BASE COUNT 51878 a 43372 c 44058 g 47545 t 600 others  
 ORIGIN  
 Query Match 100.0%; Score 21; DB 2; length 187453;  
 Best Local Similarity 100.0%; Pred. NO. 9.3;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 cagcagcagagctccatcat 21  
 |||  
 Db 124875 CAGCAGCAGAGCTCCTCATCATC 124855  
 -----  
 RESULT 12  
 AF311103/C 234431 bp DNA linear PRI 06-FEB-2002  
 LOCUS  
 DEFINITION Homo sapiens chromosome 8 clone scb-212e3 map 8p12, complete  
 sequence.  
 ACCESSION AF311103 GI:18542957  
 VERSION  
 KEYWORDS HTG; HTGS; FULLTOP; HTGS\_ACTIVEFIN.  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 234431)  
 AUTHORS Wen, G., Blechschmidt, K., Baumgart, C., Taudin, S., Baumgart, C. and  
 Platzer, M.  
 JOURNAL Chromosome 8 genomic sequence  
 TITLE Unpublished  
 COMMENT 2 (bases 1 to 234431)  
 Blechschmidt, K., Jandrig, B., Baumgart, C., Dette, M.D., Jahn, N.,  
 Menzel, U., Schilhabel, M.B., Wen, G., Taudien, S. and Rosenthal, A.  
 Direct Submission  
 Submitted (04-OCT-2000) Genome Analysis, Institute of Molecular  
 Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
 3 (bases 1 to 234431)  
 Wen, G. and Platzer, M.  
 JOURNAL Direct Submission  
 Submitted (06-FEB-2002) Genome Analysis, Institute of Molecular  
 Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany  
 On Feb 6, 2002 this sequence version replaced gi:14389409.  
 -----  
 Genome Center

Center: Institute of Molecular Biotechnology  
 Center code: IMB  
 Web site: <http://genome.imb-jena.de/>  
 Contact: [gscj-submit@genome.imb-jena.de](mailto:gscj-submit@genome.imb-jena.de)  
 ----- Project Information -----  
 Center project name: H319  
 Center clone name: SCB-212e3  
 ----- Summary Statistics -----  
 Sequencing vector: M13; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 233404 bases at least Q40  
 Consensus quality: 234413 bases at least Q30  
 Consensus quality: 234428 bases at least Q20  
 Quality coverage: 5.81 x in Q20 bases; sum-of-coverage

----- Sequence Quality Assessment -----  
 This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

## FEATURES

## source

1..234431  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8p12"  
 /clone="SCB-212e3"

BASE COUNT 62993 a 56257 c 55890 g 59291 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 234431;  
 Best Local Similarity 100.0%; Pred. No. 9.2;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 cagcagcagagcttcattcat 21  
 |||||||  
 Db 13996 CAGCAGCAGAGCTTCATCAT 13976

RESULT 13  
 DOG6P80 1636 bp mRNA linear MAM 27-APR-1993  
 LOCUS dog glycoprotein 80 (gp80) mRNA, complete cds.  
 DEFINITION M55251.M38757  
 VERSION M55251.1 GI:163953  
 KEYWORDS glycoprotein 80.  
 SOURCE Canis familiaris kidney cDNA to mRNA.  
 ORGANISM Canis familiaris  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 1636)  
 Hartmann,K., Rauch,U., Urban,J., Parczyk,K., Diel,P., Pilarsky,C., Appel,D., Haase,W., Mann,K., Weller,A. and Koch-Brandt,C.  
 Molecular cloning of gp 80, a glycoprotein complex secreted by kidney cells in vitro and in vivo: A link to the reproductive system and to the complement cascade  
 J. Biol. Chem. 266, 9924-9931 (1991)  
 91236776

## JOURNAL MEDLINE FEATURES

## source

1..1636  
 /organism="Canis familiaris"  
 /db\_xref="taxon:9615"  
 /cell\_line="MDCK"  
 /cell\_type="epithelial-like"  
 /tissue\_type="kidney"  
 1..1636  
 gene

misc\_feature  
 54..119  
 /gene="gp80"  
 /note="leader peptide"  
 54..1391  
 /gene="gp80"  
 /codon\_start=-1  
 /evidence=experimental  
 /product="glycoprotein 80"  
 /protein\_id="AAA30846.1"  
 /db\_xref="GI:163954"

## misc\_feature

120..732  
 /gene="gp80"

misc\_feature  
 733..1388  
 /gene="gp80"  
 /note="p 45 subunit"

## polyA\_signal

polyA\_signal  
 1607..1612  
 /evidence=experimental  
 /note="p 1612"

BASE COUNT 439 a 440 c 405 g 352 t  
 ORIGIN

Query Match 92.4%; Score 19.4; DB 4; Length 1636;  
 Best Local Similarity 95.2%; Pred. No. 64;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 cagcagcagagcttcattcat 21  
 |||||||  
 Db 74 CAGCAGCAGAGCTTCATCAT 54

RESULT 14  
 AF118852/c 1678 bp mRNA linear MAM 19-NOV-2001  
 LOCUS AF118852  
 DEFINITION Oryctolagus cuniculus apolipoprotein J/clusterin mRNA, complete cds.  
 ACCESSION AF118852  
 VERSION AF118852.1 GI:4580642  
 KEYWORDS rabbit.  
 SOURCE Oryctolagus cuniculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

REFERENCE 1 (bases 1 to 1678)  
 Miyata,M., Biro,S., Kaieda,H., Eto,H., Orihara,K., Kihara,T., Ohta,H., Matsushita,N., Matsuyama,T. and Tel.C.  
 Apolipoprotein J/clusterin is induced in vascular smooth muscle cells after vascular injury  
 Circulation 104 (12), 1407-1412 (2001)  
 21444720

## JOURNAL MEDLINE FEATURES

## source

2 (bases 1 to 1678)  
 Miyata,M., Biro,S., Kaieda,H., Eto,H., Orihara,K., Kihara,T., Ohta,H., Matsushita,N., Matsuyama,T. and Tel.C.  
 Direct Submission  
 Submitted (08-JAN-1999) First Department of Internal Medicine, Kagoshima University, 8-35-1 Sakuragaoka, Kagoshima 890-8520, Japan  
 Location/Qualifiers

source 1. .1678  
/organism="Oryctolagus cuniculus"  
/strain="Japanese white"  
/db\_xref="taxon:9986"  
69.1412  
/codon\_start=1  
/product="apolipoprotein J/clusterin"  
/protein\_id="A062461.1"  
/db\_xref="GI:4580643"  
/translation="MTLLGCVGLLSMERGOVLGDOLVSDNELQEMSTQSKYIDRE  
IONAVGVOEITKLEKTEPERKTLISYLEAKKKEDALNETRDSERKKAPEVGN  
ETMALHEBCKPCLKOTCKFYARVCBSGSLVQLEPLNQSPPFYINCRIDS  
LLENDKQSHLDVWQDSFNKATGIMDELFDQRFTHPQDTFTSPSTFRPLHT  
AKSRLVNTIMPUSLYGPLNFDMPQFPEFMHQAOQAMDVHLSPAQTPNVEITGG  
PDRRAVCKEIRHNSGTCLRMKDOCAKCOEILSDVCSANNPSQNDLRQELNDSLRAE  
LTKRYNELQSYQMKMLNTSSLDQPNQEFNMVSOIANLTQGPQYLYLRVSTVSHTS  
ESEAPSVFVYVYKLFSDSPITITIPPEEVSNDNPFMEVVAEKALOEYRKKRYE"

BASE COUNT 400 a 523 c 452 g 303 t

ORIGIN

Query Match 92.4%; Score 19.4; DB 4; Length 1678;  
Best Local Similarity 95.2%; Pred. No. 64;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 cagcagcagagcttcacat 21  
|||||

Db 86 CAGCAGCAGAGCTTCATCTT 66

RESULT 15  
AF063235 3287 bp mRNA linear PLN 22-AUG-2000  
LOCUS Pneumocystis carinii clone 15 variant 2 major surface glycoprotein  
DEFINITION mRNA, complete cds.  
ACCESSION AF063235  
VERSION AF063235.1 GI:4321810  
KEYWORDS  
SOURCE  
ORGANISM  
Pneumocystis carinii.  
Pneumocystis carinii  
Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;  
Pneumocystidaceae; Pneumocystis.  
REFERENCE 1 (bases 1 to 3287)  
AUTHORS Huang,S.N., Angus,C.W., Turner,R.E., Sorial,V. and Kovacs,J.A.  
TITLE Identification and characterization of novel variant major surface  
glycoprotein gene families in rat Pneumocystis carinii  
J. Infect. Dis. 179 (1), 192-200 (1999)  
JOURNAL 99059868  
MEDLINE 9841839  
PUBMED  
REFERENCE 2 (bases 1 to 3287)  
AUTHORS Huang,S.N., Angus,C.W., Turner,R.E., Sorial,V. and Kovacs,J.A.  
TITLE Direct Submission  
JOURNAL Submitted (05-MAY-1998) Critical Care Medicine Department, Clinical  
Center, National Institutes of Health, 9000 Rockville Pike,  
Bethesda, MD 20892, USA  
FEATURES  
Source  
Location/Qualifiers  
1..3287  
/organism="Pneumocystis carinii"  
/specific\_host="rat"  
/db\_xref="taxon:4754"  
/clone="15"  
41..3175  
/note="v2MSG"  
/codon\_start=1  
/product="variant 2 major surface glycoprotein"  
/db\_xref="GI:4321811"  
/translation="MFEISMINKIIVGLILVPIYGNQNTLRYRDSVSAENMLPE  
DEVYKEDIYAVILGDEYVEYKCKSLLEECNELVIDPELKNVHTKKEICDNIE  
OKCAGLNDKTEVEFSNPNRIYDALGNLKDQCKNNEKCIETLEAPRHDPKENCYK  
LRECYKIKROKADDELRLALGNKRENDKCKEKTICPILSRSDPEMFSCLNT  
KTCDDLSKLGSLPLKREIKDKDLKCKEYERLEKCHFTGEACDDAKQVFEBCCK  
NNTYKAPESDFSPKPRASLTLMIIGLEDVYNAEKDILIGKAGVDPRKLGITVPLQ

DILLSDQVNMEDTVKKNIMLERCNDLKHLGNEMTKLCEDGKKNCKILLDVNVK  
ERCTFKSALTEFOKDLSTQVNDDEKSELPMGHLPTFTGECTELVSECFYLKNACTN  
KIDACQNVHACYKIKGRDLINKFPKELRGKLGDRNSLSLKKCOEYVENCETKL  
DERYLPCLYPRKLCYVLSNDIYVQSRFGLVILDGDRSLERKHLKEIGEKDELD  
SNSNRKCATIKILCEYFVRAEKFRTYELKREDDALDYDQNTKALHEKCDTLRGEN  
PFSICALPRETCRYVYHTRQDCRYLKDMMKNEGILEKTEOKKANANETLVBELCT  
TWGRCHQLVNCPDQJGKERNGDQCEKLDKCRDPTKRLKSDDELTHLKGSLN  
DKCRDALRKHCPDLQNGTFRILIDNCEQNTYKELCEKLVRYQNRKCPYLKTDLERAK  
DELTEKDEYKAKOAESTEAARLLLSRONSDEQVSPLSAESSGSSSELPP  
PPPPONGPPTPDGASGTPGSGMPNRYKFGLYRREYVAGVSEAEVAKAFDATYIALEL  
YLELKEEKALELDGCFRDCPPNSKEVCEIDTLCKEIEPLKDKPYEKITTEPCTLLQ  
TDIWTSTVSTSTVSTSMRCKPRTCTDSDSKETQKQEBEVEYKPNQMKIRVPM  
IKIMLGTVVGM"

BASE COUNT 1277 a 407 c 731 g 872 t

ORIGIN

Query Match 87.6%; Score 18.4; DB 8; Length 3287;  
Best Local Similarity 95.0%; Pred. No. 2e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 cagcagcagagcttcacat 20  
|||||

Db 2565 CAGCAGCAGAGCTTCATCA 2584

Search completed: May 17, 2002, 16:04:13  
Job time: 16113 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 15:25:36 ; Search time 3628.41 Seconds  
(without alignments)  
78.116 Million cell updates/sec

Title: US-09-913-325-5

Perfect score: 21  
Sequence: 1 attgtctgagacgtctgctgctc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_estl2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pin:\*  
16: em\_gss\_vtl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	152	9	AM161759
2	21	100.0	156	10	BF922862
3	21	100.0	161	10	BF921576
4	21	100.0	165	10	BF92381
5	21	100.0	167	10	BF929953
6	21	100.0	169	10	BE766895
7	21	100.0	177	9	AL048592
8	21	100.0	177	10	BF958041
9	21	100.0	184	10	BF958930
10	21	100.0	185	10	BF036860
11	21	100.0	190	10	BT037432
12	21	100.0	193	10	BF944443
13	21	100.0	196	9	AT1745406
14	21	100.0	201	10	BF957666
15	21	100.0	207	10	BF957858
16	21	100.0	218	9	AM807875
17	21	100.0	220	9	AT904991

18	21	100.0	223	10	BF920102	BF920102
19	21	100.0	224	9	AM161224	AM161224
20	21	100.0	226	10	BF935119	BF935119
21	21	100.0	227	10	BE841947	BE841947
22	21	100.0	229	10	BF950916	BF950916
23	21	100.0	236	9	AA339299	AA339299
24	21	100.0	238	10	BF921875	BF921875
25	21	100.0	243	10	BF923639	BF923639
26	21	100.0	243	10	AA336628	AA336628
27	21	100.0	244	10	BE766870	BE766870
28	21	100.0	248	10	BF923643	BF923643
29	21	100.0	252	10	BF923633	BF923633
30	21	100.0	253	10	BE887875	BE887875
31	21	100.0	253	10	BI041998	BI041998
32	21	100.0	256	10	BI041248	BI041248
33	21	100.0	257	9	AA365743	AA365743
34	21	100.0	262	10	BF947155	BF947155
35	21	100.0	266	9	AA352030	AA352030
36	21	100.0	268	10	BF920149	BF920149
37	21	100.0	272	10	BF956582	BF956582
38	21	100.0	273	10	BF956584	BF956584
39	21	100.0	274	10	BI036852	BI036852
40	21	100.0	277	9	AA351754	AA351754
41	21	100.0	280	10	BF935974	BF935974
42	21	100.0	283	10	BF947214	BF947214
43	21	100.0	284	10	BC995096	BC995096
44	21	100.0	293	9	AA318247	AA318247
45	21	100.0	294	9	AA498619	AA498619

#### ALIGNMENTS

RESULT 1  
LOCUS AM161759  
DEFINITION au70a10.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone  
IMAGE:2781594 3' similar to gb:xl4723 CLUSTERIN PRECURSOR (HUMAN);  
mRNA sequence.  
AM161759  
AM161759.1 GI:6300792

ACCESSION AM161759  
VERSION  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Maira,M., Martin  
J., Moore,B., Schellenberg,K., Stepien,M., Tan,F., Thelsting,B.,  
White,Y., Wylie,T., Waterston,R. and Wilson,R.

WashU-NCI human EST Project  
Unpublished (1997)  
Other ESTs: au70a10.y1  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available royalty-free through LINT; contact the  
Image Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40UP from Gibco.

#### FEATURES

source  
1..152  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2781594"  
/clone\_lib="Schneider fetal brain 00004"  
/sex="male"  
/tissue\_type="frontal lobe"  
/dev\_stage="5 months post-conception"  
/lab\_host="DH10B"

/note="Organ: brain; Vector: pBluescript SK (Stratagene); Site\_1: SctI; Site\_2: XhoI; Double-stranded cDNA was prepared from human fetal brain tissue. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-GAGAGAGAGAGAGAGAGAGATCTTAATTAAATTATATCCCCCCCC-3' and 3' adaptor sequence: 5'-GAGAGAGAGAGAGAGAGAGATCTTAATTAAATTATATCCCCCCCC-3' size-selected for >0.5 kb inserts and has an average insert size estimated at 1.2 kb. This library was constructed using the CAP-trapper method for full-length enrichment and has not undergone amplification. Library was constructed by Dr. Claudio Schneider (LNCIB-Area Science Park, Trieste, Italy)."

BASE COUNT 27 a 43 c 28 g 54 t

ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 152;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgtctgagacgcgtctgctc 21  
|||||  
Db 100 ATTGCTGAGACGCTGCTGTC 120

RESULT 2  
LOCUS BF922862 156 bp mRNA linear EST 19-JAN-2001  
DEFINITION QV4-NT0247-201100-262-g12 NT0247 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF922862  
VERSION BF922862.1 GI:12318659  
KEYWORDS EST  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 156)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl-QV4&c2-QV4-NT0247-201100-262-g12&t3-2000-11-20&t4-1)  
Seq primer: puc 18 forward  
High quality sequence start: 15.  
High quality sequence stop: 155.  
Location/Qualifiers

## FEATURES

1..156  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NT0247"  
/dev\_stage="Adult"  
/note="Organ: nervous\_tumor; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. letters patent application

No. 196/716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 38 a 30 c 51 g 37 t

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 156;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atgtctgagacgcgtctgctc 21  
|||||  
Db 94 ATTGCTGAGACGCTGCTGTC 74

RESULT 3  
LOCUS BF921576 161 bp mRNA linear EST 19-JAN-2001  
DEFINITION MR1-NT0179-171100-004-g01 NT0179 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF921576  
VERSION BF921576.1 GI:12317464  
KEYWORDS EST  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 161)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl-MR1&c2-MR1-NT0179-171100-004-g01&t3-2000-11-17&t4-1)  
Seq primer: puc 18 forward  
High quality sequence start: 16  
High quality sequence stop: 161.  
Location/Qualifiers

## FEATURES

1..161  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NT0179"  
/dev\_stage="Adult"  
/note="Organ: nervous\_tumor; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. letters patent application No. 196/716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## BASE COUNT

44 a 29 c 53 g 35 t

## ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 161;  
Best Local Similarity 100.0%; Pred. No. 4.5;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagacgctctgctc 21  
|||||

Db 101 ATTGTCTGAGACCGCTCTGCTC 81

RESULT 4  
BF929381 165 bp mRNA linear EST 19-JAN-2001  
LOCUS IL2-NT0198-081200-280-G01 NT0198 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BF929381  
ACCESSION  
VERSION BF929381.1 GI:12327509  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 165)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL2&tl2=IL2-NT0198-  
081200-280-G01&tl3=2000-12-08&tl4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 165.  
FEATURES  
Location/Qualifiers  
source  
1..165  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NT0198"  
/dev\_stage="Adult"  
/note="Organ: nervous.tumor; Vector: puc18; Site.1: Sma1;  
Site.2: Sma1; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

BASE COUNT 27 a 47 c 31 g 60 t

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 165;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagacgctctgctc 21  
|||||

Db 112 ATTGTCTGAGACCGCTCTGCTC 132

RESULT 5  
BF929953 167 bp mRNA linear EST 19-JAN-2001  
LOCUS IL2-NT0198-111200-293-A03\_1 NT0198 Homo sapiens cDNA, mRNA  
DEFINITION

sequence.  
BF929953  
KEYWORDS BF929953.1 GI:12328185  
EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 167)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL2&tl2=IL2-NT0198-  
111200-293-A03\_1&tl3=2000-12-11&tl4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 165.  
FEATURES  
Location/Qualifiers  
source  
1..167  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NT0198"  
/dev\_stage="Adult"  
/note="Organ: nervous.tumor; Vector: puc18; Site.1: Sma1;  
Site.2: Sma1; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

BASE COUNT 60 a 32 c 48 g 27 t

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 167;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagacgctctgctc 21  
|||||

Db 54 ATTGTCTGAGACCGCTCTGCTC 34

RESULT 6  
BE766895 169 bp mRNA linear EST 19-SEP-2000  
LOCUS RC2-NT0110-050600-013-f03 NT0110 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BE766895  
ACCESSION BE766895  
VERSION BE766895.1 GI:10196819  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 169)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=6&l2=RC2-NT0110-050  
600-013-f03&l3=2000-06-05&l4=1)

Seq primer: puc 18 forward

High quality sequence start: 18

High quality sequence stop: 169.

## FEATURES

source

1. 169

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_1lb="NT0110"

/dev\_stage="Adult"

/note="Organ: nervous\_tumor; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT

37 a 36 c 37 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 21; DB 10; Length 169;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 attgtctgagaccgtctgctc 21

|||||

Db 125 ATTGCTGTGAGACCGTCTGCTC 105

RESULT 7

AL048592/c 177 bp mRNA linear EST 01-MAR-2000

LOCUS

DEFINITION

DKFZP586H092.r1 586 (synonym: hute1) Homo sapiens cDNA clone

ACCESSION

AL048592

VERSION

AL048592.1 GI:4729143

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 177)

Poustka, A., Kleih, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

EST (Poustka, et al.)

Unpublished (1999)

Contact: Poustka A.J.

Department Lehnach

Max-Planck-Institute for Molecular Genetics

Imnestrasse 73, 14195 Berlin, Germany

Tel: +49-30-84131623

Fax: +49-30-84131128

Email: poustka@mpg-berlin-dahlem.mpg.de

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ). Email: s.wiemann@dkfz-heidelberg.de; Sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

No sl sequence available.

This clone (DKFZP586H092) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

source

1. 177

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="DKFZP586H092"

/clone\_1lb="586 (synonym: hute1)"

/tissue\_type="uterus"

/dev\_stage="adult"

/lab\_host="DH10B"

/note="Vector: pSPori1; Site\_1: NotI; Site\_2: SalI/MluI"

BASE COUNT

40 a 41 c 64 g 32 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 21; DB 9; Length 177;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 attgtctgagaccgtctgctc 21

|||||

Db 143 ATTGCTGTGAGACCGTCTGCTC 123

RESULT 8

BF958041/c 177 bp mRNA linear EST 22-JAN-2001

LOCUS

DEFINITION

RC5-NN1165-281100-033-E03 NN1165 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BF958041

VERSION

BF958041.1 GI:12375316

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 177)

Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W., Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=RC5&l2=RC5-NN1165-281100-033-E03&l3=2000-11-28&l4=1)

Seq primer: puc 18 forward

High quality sequence stop: 177.

## FEATURES

source

1. 177

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_1lb="NN1165"

```
/dev stage="Adult"
/Note="Organ: nervous normal; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      63 a      34 c      52 g      28 t
ORIGIN
Query Match      100.0%; Score 21; DB 10; Length 177;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 atgtctgagaccgtctgctc 21
        |||
        68 ATTGCTCGAGACCGCTCTGCTC 48

RESULT 9
BF958930/c      184 bp      mRNA      linear      EST 22-JAN-2001
LOCUS      PM1-NNI200-011200-009-g08 NNI200 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BF958930
ACCESSION      BF958930
VERSION      EST.
KEYWORDS      GI:12376205
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 184)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1&t2=PM1-NNI200-
011200-009-g08&t3=2000-12-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 184.
FEATURES
Location/Qualifiers
1..184
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="NNI200"
/dev_stage="Adult"
/Note="Organ: nervous normal; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      47 a      36 c      61 g      40 t
ORIGIN
```

```
Query Match      100.0%; Score 21; DB 10; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 atgtctgagaccgtctgctc 21
        |||
        108 ATTGCTCGAGACCGCTCTGCTC 88

RESULT 10
BI036860/c      185 bp      mRNA      linear      EST 14-JUN-2001
LOCUS      MR4-NT0140-080101-209-c09 NT0140 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BI036860
ACCESSION      BI036860
VERSION      EST.
KEYWORDS      GI:14443486
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 185)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR4&t2=MR4-NT0140-
080101-209-c09&t3=2001-01-08&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 134.
FEATURES
Location/Qualifiers
1..185
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="NT0140"
/dev_stage="Adult"
/Note="Organ: nervous tumor; Vector: puc18; Site:1: SmaI;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      41 a      42 c      62 g      40 t
ORIGIN
Query Match      100.0%; Score 21; DB 10; Length 185;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 atgtctgagaccgtctgctc 21
        |||
        109 ATTGCTCGAGACCGCTCTGCTC 89
```

RESULT 11  
 B1037432/c 190 bp mRNA linear EST 14-JUN-2001  
 LOCUS B1037432  
 DEFINITION CM3-NT0267-110101-588-h12 NT0267 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION B1037432  
 VERSION B1037432.1 GI:14444058  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 190)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 COMMENT Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&f2=CM3-NT0267-110101-588-112&f3=2001-01-11&f4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 4  
 Location/Qualifiers  
 1. 190  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_id="NT0267"  
 /dev\_stage="Adult"  
 /note="Organ: nervous.tumor; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
 BASE COUNT 76 a 38 c 49 g 27 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 4.7;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagacgctcgtgc 21  
 ||||||||||||||||||||  
 Db 36 ATTGCTGAGACCGCTGCTC 16

RESULT 12  
 BF944445/c 193 bp mRNA linear EST 22-JAN-2001  
 LOCUS BF944445  
 DEFINITION RC5-NM1165-131000-021-D04 NM1165 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF944445  
 VERSION BF944445.1 GI:12361720  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 193)  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 20202663  
 COMMENT Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&f2=RC5-NM1165-131000-021-D04&f3=2000-10-13&f4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 193.  
 Location/Qualifiers  
 1. 193  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_id="NM1165"  
 /dev\_stage="Adult"  
 /note="Organ: nervous.normal; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
 BASE COUNT 70 a 39 c 57 g 27 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 4.7;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagacgctcgtgc 21  
 ||||||||||||||||||||  
 Db 72 ATTGCTGAGACCGCTGCTC 52

RESULT 13  
 A1745406 196 bp mRNA linear EST 17-DEC-1999  
 LOCUS A1745406  
 DEFINITION wc37d01.x1 NCI-CCAP\_Pt28 Homo sapiens cDNA clone IMAGE:2320801 3'  
 ACCESSION A1745406  
 VERSION A1745406.1 GI:5113694  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 196)  
 AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ccgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CCAP), Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: ccaps-remail.nih.gov  
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LML at:  
[www-bio.lml.gov/bhrp/image/image.html](http://www-bio.lml.gov/bhrp/image/image.html)  
 Insert length: 264 Std Error: 0.00  
 Seq primer: -40UP from Gibco.

## FEATURES

source

Location/Qualifiers  
 1. 196  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2320801"  
 /clone\_lib="NCI-CGAP\_Pr28"  
 /sex="male"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)  
 with a modified polylinker; Plasmid DNA from the  
 normalized library NCI-CGAP\_Pr22 was prepared, and ss  
 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (clonesids  
 985608-986759, 1101192-1101959, and 121928-1220615)."  
 Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 39 a 60 c 35 g 62 t

ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 196;  
 Best Local Similarity 100.0%; Pred. No. 4.7;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagacgctctgctc 21  
 |||||||||||||||||||||  
 Db 101 ATTGCTGAGACGCTCTGCTC 121

RESULT 14  
 BE957666/c 201 bp mRNA linear EST 22-JAN-2001  
 LOCUS PM1-NN1200-251100-007-c09 NN1200 Homo sapiens CDNA, mRNA sequence.  
 DEFINITION BE957666  
 ACCESSION BE957666.1 GI:12374941  
 VERSION  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 201)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunsfeld, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
 , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MERLIN 20202663  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1&lt2=PM1-NN1200->

251100-007-c09&lt3=2000-11-25&lt4-1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 14  
 High quality sequence stop: 201.  
 Location/Qualifiers

## FEATURES

source

1. 201  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NN1200"  
 /dev\_stage="Adult"  
 /note="Organ: nervous normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 52 a 39 c 71 g 39 t

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 4.8;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagacgctctgctc 21  
 |||||||||||||||||||||  
 Db 125 ATTGCTGAGACGCTCTGCTC 105

RESULT 15  
 BE957858/c 207 bp mRNA linear EST 22-JAN-2001  
 LOCUS PM1-NN1200-271100-008-a07 NN1200 Homo sapiens CDNA, mRNA sequence.  
 DEFINITION BE957858  
 ACCESSION BE957858.1 GI:12375133  
 VERSION  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 207)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunsfeld, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
 , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MERLIN 20202663  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1&lt2=PM1-NN1200->

271100-008-a07&lt3=2000-11-27&lt4-1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 24  
 High quality sequence stop: 187.  
 Location/Qualifiers

## FEATURES

source

1. 207  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NN1200"  
 /dev\_stage="Adult"

/note="Organ: nervous\_normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."  
 BASE COUNT 51 a 45 c 71 g 40 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 4.8;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 attgtctgagaccgtctgtgc 21  
 ||||||||||||||||||||  
 Db 143 ATTGCTGAGACCGCTGTGTC 123

Search completed: May 17, 2002, 15:25:38  
 Job time: 14515 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 17:58:43 ; Search time 147.64 Seconds  
(without alignments)  
66.583 Million cell updates/sec

Title: US-09-913-325-5

Perfect score: 21

Sequence: 1 atgtctgagaccgtctgctc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 424068 seqs, 234053524 residues

Total number of hits satisfying chosen parameters: 848136

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_New:\*  
1: /cgn2\_6/ptodata/1/pna/PCR\_NEW\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:\*  
7: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21	100.0	1614	6	US-10-119-428-31
C 2	21	100.0	1839	7	US-60-365-384-115
C 3	15.8	75.2	355	5	US-10-106-698-1758
C 4	15.2	72.4	200	5	US-09-539-331D-26988
C 5	15.2	72.4	251	5	US-09-975-254-18048
C 6	15	71.4	302	5	US-09-539-331D-33835
C 7	14.8	70.5	1482	6	US-10-103-627-30
C 8	14.8	70.5	4109	5	US-09-741-148A-1
C 9	14.8	70.5	23668	5	US-09-741-148A-3
C 10	14.8	70.5	33958	6	US-10-105-299-14037
C 11	14.6	69.5	932	6	US-10-105-299-1118
C 12	14.6	69.5	1644	5	US-09-053-375B-823
C 13	14.6	69.5	6162	6	US-10-105-299-12825
C 14	14.4	68.6	252	5	US-09-540-210B-2398
C 15	14.4	68.6	494	6	US-10-113-872-45
C 16	14.4	68.6	518	5	US-09-539-331D-31816
C 17	14.4	68.6	1611	6	US-10-105-299-14778
C 18	14.4	68.6	1611	6	US-10-105-299-15187
C 19	14.2	67.6	331	5	US-09-673-476-310
C 20	14.2	67.6	472	5	US-09-539-331D-33623
C 21	14.2	67.6	510	5	US-09-539-331D-34970
C 22	14.2	67.6	511	6	US-10-124-805-443
C 23	14.2	67.6	614	6	US-10-105-299-11648
C 24	14.2	67.6	628	6	US-10-105-299-2095
C 25	14.2	67.6	648	5	US-09-930-213-190
C 26	14.2	67.6	1326	5	US-09-749-728B-12

27	14.2	67.6	1428	5	US-09-216-387-2	Sequence 2, Appli
28	14.2	67.6	2177	5	US-09-722-377-4	Sequence 4, Appli
C 29	14.2	67.6	2248	5	PCT-US02-09944-231	Sequence 231, App
C 30	14.2	67.6	2502	6	US-10-105-299-1215	Sequence 1215, App
C 31	14.2	67.6	2764	5	US-09-978-403A-258	Sequence 258, App
C 32	14.2	67.6	2764	5	US-09-978-544A-258	Sequence 258, App
C 33	14.2	67.6	2764	5	US-09-978-681A-258	Sequence 258, App
C 34	14.2	67.6	2764	5	US-09-978-757A-258	Sequence 258, App
C 35	14.2	67.6	2764	5	US-09-978-564A-258	Sequence 258, App
C 36	14.2	67.6	2764	5	US-09-999-831A-258	Sequence 258, App
C 37	14.2	67.6	2764	5	US-09-999-831A-258	Sequence 258, App
C 38	14.2	67.6	2764	5	US-09-978-375A-258	Sequence 258, App
C 39	14.2	67.6	2764	6	US-10-013-921A-258	Sequence 258, App
C 40	14.2	67.6	2764	6	US-10-013-921A-258	Sequence 258, App
C 41	14.2	67.6	2764	6	US-10-013-918A-258	Sequence 258, App
C 42	14.2	67.6	2764	6	US-10-017-082A-258	Sequence 258, App
C 43	14.2	67.6	2764	6	US-10-121-062-117	Sequence 117, App
C 44	14.2	67.6	4282	5	US-09-053-375B-1069	Sequence 1069, App
C 45	14.2	67.6	4997	6	US-10-105-299-9206	Sequence 9206, App

## ALIGNMENTS

RESULT 1  
US-10-119-428-31/c  
Sequence 31, Application US/10119428  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Xu, Chongjun  
APPLICANT: Wehrman, Tom  
APPLICANT: Ren, Feiyan  
APPLICANT: Ma, Yundong  
APPLICANT: Zhou, Ping A.  
APPLICANT: Zhao, Qing A.  
APPLICANT: Dmanac, Radolje T.  
TITLE OF INVENTION: Novel Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 789CIP2  
CURRENT APPLICATION NUMBER: US/10/119,428  
CURRENT FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: 09/596,193  
PRIOR FILING DATE: 2000-06-17  
PRIOR APPLICATION NUMBER: 09/574,454  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: 09/519,705  
PRIOR FILING DATE: 2000-03-07  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: PL\_FL\_genes Version 1.0  
SEQ ID NO 31  
LENGTH: 1614  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (116)..(1465)  
US-10-119-428-31

Query Match 100.0%; Score 21; DB 6; Length 1614;  
Best local Similarity 100.0%; Pred. No. 0.06;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgtctgagaccgtctgctc 21  
DB 202 ATGTCTGAGACCGCTCTGCTC 182  
RESULT 2  
US-60-365-384-115/c

```

Sequence 115: Application US/60365384
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Wang, Jian-rui
APPLICANT: Wang, Zhiwei
APPLICANT: Zhang, Jie
APPLICANT: Zhou, Ping
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Ghosh, Malabika
APPLICANT: Zhao, Qing A.
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyang
APPLICANT: Xue, Aidong J.
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Gezhi
TITLE OF INVENTION: Novel Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 814
CURRENT APPLICATION NUMBER: US/60/365,384
CURRENT FILING DATE: 2002-03-14
NUMBER OF SEQ ID NOS: 666
SOFTWARE: pc_FL_genes Version 6.0
SEQ ID NO 115
LENGTH: 1839
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (132)..(4316)
US-60-365-384-115

```

	Query Match	100.0%	Score 21:	DB 7;	Length 1839;
	Best Similarity	100.0%	Pred. NO.	0.061,	
	Matches 21: Conservative	0;	Mismatches	0;	Indels 0;
QY	1 atgtctgagacgcgtcgtgc	21			
db	218 ATTGCTGTAGACCGTCTGTGC	198			

```

RESULT 3
US-10-106-698-1758/c
: Sequence 1758, Application US/10106698
: GENERAL INFORMATION:
: APPLICANT: Ruben et al.
: TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
: FILE REFERENCE: PA005F1
: CURRENT APPLICATION NUMBER: US/10/106, 698
: CURRENT FILING DATE: 2002-03-27
: PRIOR APPLICATION NUMBER: PCT/US00/26524
: PRIOR FILING DATE: 2000-09-28
: PRIOR APPLICATION NUMBER: US 60/157,137
: PRIOR FILING DATE: 1999-09-29
: PRIOR APPLICATION NUMBER: US 60/163,280
: PRIOR FILING DATE: 1999-11-03
: NUMBER OF SEQ ID NOS: 8564
: SOFTWARE: PatentIn Ver. 3.0
: SEQ ID NO 1758
: LENGTH: 355
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (8)..(8)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc_feature
: LOCATION: (353)..(353)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc_feature
: LOCATION: (355)..(355)

```

OTHER INFORMATION: n equals a,t,g, or c  
US-10-106-698-1758

Query Match	75.2%	Score 15.8;	DB 6;	length 355;
Best Local Similarity	89.5%	Pred. No. 30;		
Matches 17; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

```

QY      3  tgtctgagacgctctgtc  21
          |||||  || |||||
Db      308  TGTCTGAGTCCCTCTGTC  290

```

```

RESULT      4
US-09-539-33ID-26988/C
; Sequence 26988, Application US/0953933ID
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mulahy, Sara J.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CJP
; CURRENT APPLICATION NUMBER: US/09/539,33ID
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 26988
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00420089
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 106
; OTHER INFORMATION: a, t, c, g, or other
US-09-539-33ID-26988

```

```

Query Match          72.4%      Score 15.2:  DB 5;  Length 200;
Best Local Similarity 85.0%      Pred. No. 59;
Matches 17;  Conservative 0;  Mismatches 3;  Indels 0;  Gaps 0;

Oy      2  ttgtctgagacctctgtgc 21
          |||||  ||  |||||  ||
Db      26  ttgtctgacacctgtctgtc 7

RESULT      5
US-09-975-254-18048/c
; Sequence 18048, Application US/09975254
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15309)B
; CURRENT APPLICATION NUMBER: US/09/975,254
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US/09/263,191
; PRIOR FILING DATE: 1999-03-05
; NUMBER OF SEQ ID NOS: 31255
; SEQ ID NO 18048
;
; TYPE: DNA
; LENGTH: 251
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 7009955068H1
;

```

US-09-975-254-18048

Query Match  
Best Local Similarity 72.4%; Score 15.2; DB 5; Length 251;  
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 ttgtctgagaccgtctgctg 21  
DB 226 ttgtctgagaccgtctgctg 207

RESULT 6  
US-09-539-331D-33835/c  
Sequence 33835, Application US/09539331D

GENERAL INFORMATION:  
APPLICANT: Seilhamer, Jeffrey J.  
APPLICANT: Deleage, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mulhaly, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE  
FILE REFERENCE: PD-1022 CIP  
CURRENT APPLICATION NUMBER: US/09/539,331D  
CURRENT FILING DATE: 2000-03-30  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 40961  
SOFTWARE: PERL Program  
SEQ ID NO 33835  
LENGTH: 302  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No: hu01211390  
US-09-539-331D-33835

Query Match  
Best Local Similarity 71.4%; Score 15; DB 5; Length 302;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 tctgagaccgtctg 19  
DB 284 tctgagaccgtctg 270

RESULT 7

US-10-102-627-30/c  
Sequence 30, Application US/10102627  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT24CI  
CURRENT APPLICATION NUMBER: US/10/102,627  
CURRENT FILING DATE: 2002-03-22  
NUMBER OF SEQ ID NOS: 110  
Prior Application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 30  
LENGTH: 1482  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1453)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc.feature  
LOCATION: (1474)  
OTHER INFORMATION: n equals a,t,g, or c  
US-10-102-627-30

Query Match  
Best Local Similarity 70.5%; Score 14.8; DB 6; Length 1482;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 ttgtctgagaccgtctg 19  
DB 439 ttgtctgagaccgtctg 422

RESULT 8  
US-09-741-148A-1/c

Sequence 1, Application US/09741148A  
GENERAL INFORMATION:  
APPLICANT: Chunhua YAN et al.  
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
FILE REFERENCE: CLO00566  
CURRENT APPLICATION NUMBER: US/09/741,148A  
CURRENT FILING DATE: 2002-04-15  
PRIOR APPLICATION NUMBER: 60/206,982  
PRIOR FILING DATE: 2000-05-25  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 4109  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-741-148A-1

Query Match  
Best Local Similarity 70.5%; Score 14.8; DB 5; Length 4109;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 gctctgagaccgtctgctc 21  
DB 1127 gctctgagaccgtctgctc 1110

RESULT 9  
US-09-741-148A-3/c

Sequence 3, Application US/09741148A  
GENERAL INFORMATION:  
APPLICANT: Chunhua YAN et al.  
TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
FILE REFERENCE: CLO00566  
CURRENT APPLICATION NUMBER: US/09/741,148A  
CURRENT FILING DATE: 2002-04-15  
PRIOR APPLICATION NUMBER: 60/206,982  
PRIOR FILING DATE: 2000-05-25  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 23668  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(23668)  
OTHER INFORMATION: n = A,T,C or G  
US-09-741-148A-3

Query Match  
Best Local Similarity 70.5%; Score 14.8; DB 5; Length 23668;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 gctctgagaccgtctgctc 21  
DB 1127 gctctgagaccgtctgctc 1110

Db 15242 GTCTGAGACGCTCTGTC 15225

RESULT 10

US-10-105-299-14037  
; Sequence 14037, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14037  
; LENGTH: 33958  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-14037

Query Match 70.5%; Score 14.6; DB 6; Length 33958;  
Best Local Similarity 88.9%; Pred. No. 1.5e+02;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 gtctgagagccgtctgtc 21  
Db 30348 gtctgagagagctgtc 30365

RESULT 11

US-10-105-299-1118  
; Sequence 1118, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1118  
; LENGTH: 932  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (20)-(20)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc.feature  
; LOCATION: (28)-(28)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-105-299-1118

Query Match 69.5%; Score 14.6; DB 6; Length 932;  
Best Local Similarity 81.0%; Pred. No. 1.4e+02;  
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 attgtctgagagccgtctgtc 21  
Db 854 attgtctgagagccgtctgtc 874

RESULT 12

US-09-053-375B-823/c  
; Sequence 823, Application US/09053375B  
; GENERAL INFORMATION:  
; APPLICANT: Chenchik, Alex  
; APPLICANT: Bibilashvili, Robert  
; TITLE OF INVENTION: Nucleic Acid Arrays

; FILE REFERENCE: CLON-006  
; CURRENT APPLICATION NUMBER: US/09/053,375B  
; CURRENT FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 1543  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 823  
; LENGTH: 1644  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-053-375B-823

Query Match 69.5%; Score 14.6; DB 5; Length 1644;  
Best Local Similarity 81.0%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 attgtctgagagccgtctgtc 21  
Db 147 ATTGTCTGAGAGCCCTCTGCTC 127

RESULT 13

US-10-105-299-12825  
; Sequence 12825, Application US/10105299  
; GENERAL INFORMATION:  
; APPLICANT: Rosen, et. al  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS950  
; CURRENT APPLICATION NUMBER: US/10/105,299  
; CURRENT FILING DATE: 2002-03-26  
; NUMBER OF SEQ ID NOS: 15197  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12825  
; LENGTH: 6162  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-105-299-12825

Query Match 69.5%; Score 14.6; DB 6; Length 6162;  
Best Local Similarity 81.0%; Pred. No. 1.5e+02;  
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 attgtctgagagccgtctgtc 21  
Db 4982 attgtctgagagccgtctgtc 5002

RESULT 14

US-09-540-210B-2398/c  
; Sequence 2398, Application US/09540210B  
; GENERAL INFORMATION:  
; APPLICANT: Sellheame, Jeffrey J.  
; APPLICANT: Deleageane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE  
; FILE REFERENCE: PD-1037 CIP  
; CURRENT APPLICATION NUMBER: US/09/540,210B  
; CURRENT FILING DATE: 2002-04-03  
; PRIOR APPLICATION NUMBER: 08/972,899  
; PRIOR FILING DATE: November 18, 1997  
; PRIOR APPLICATION NUMBER: 08/395,244  
; PRIOR FILING DATE: February 27, 1995  
; PRIOR APPLICATION NUMBER: 08/722,922  
; PRIOR FILING DATE: September 27, 1996  
; PRIOR APPLICATION NUMBER: 60/005,526  
; PRIOR FILING DATE: September 29, 1995  
; PRIOR APPLICATION NUMBER: 08/824,029  
; PRIOR FILING DATE: March 25, 1997

;; PRIOR APPLICATION NUMBER: 60/014,010  
;; PRIOR FILING DATE: March 25, 1996  
;; PRIOR APPLICATION NUMBER: 08/826,847  
;; PRIOR FILING DATE: April 10, 1997  
;; PRIOR APPLICATION NUMBER: 60/015,533  
;; PRIOR FILING DATE: April 10, 1996  
;; PRIOR APPLICATION NUMBER: 08/903,555  
;; PRIOR FILING DATE: July 31, 1997  
;; PRIOR APPLICATION NUMBER: 60/023,308  
;; PRIOR FILING DATE: July 31, 1996  
;; PRIOR APPLICATION NUMBER: 08/862,178  
;; PRIOR FILING DATE: May 22, 1997  
;; PRIOR APPLICATION NUMBER: 60/018,217  
;; PRIOR FILING DATE: May 23, 1996  
;; PRIOR APPLICATION NUMBER: 08/881,589  
;; PRIOR FILING DATE: June 24, 1997  
;; PRIOR APPLICATION NUMBER: 60/021,275  
;; PRIOR FILING DATE: June 25, 1996  
;; PRIOR APPLICATION NUMBER: 08/903,802  
;; PRIOR FILING DATE: July 31, 1997  
;; PRIOR APPLICATION NUMBER: 60/023,308  
;; PRIOR FILING DATE: July 31, 1996  
;; PRIOR APPLICATION NUMBER: 08/905,881  
;; PRIOR FILING DATE: August 1, 1997  
;; PRIOR APPLICATION NUMBER: 60/025,204  
;; PRIOR FILING DATE: August 1, 1996  
;; PRIOR APPLICATION NUMBER: 08/903,471  
;; PRIOR FILING DATE: July 30, 1997  
;; PRIOR APPLICATION NUMBER: 60/023,478  
;; PRIOR FILING DATE: July 31, 1996  
;; PRIOR APPLICATION NUMBER: 08/903,556  
;; PRIOR FILING DATE: July 31, 1997  
;; PRIOR APPLICATION NUMBER: 60/025,217  
;; PRIOR FILING DATE: August 22, 1996  
;; PRIOR APPLICATION NUMBER: 08/937,142  
;; PRIOR FILING DATE: September 23, 1997  
;; PRIOR APPLICATION NUMBER: 60/026,598  
;; PRIOR FILING DATE: September 24, 1996  
;; PRIOR APPLICATION NUMBER: 08/960,746  
;; PRIOR FILING DATE: October 29, 1997  
;; PRIOR APPLICATION NUMBER: 60/030,144  
;; PRIOR FILING DATE: October 30, 1996  
;; PRIOR APPLICATION NUMBER: 08/826,847  
;; PRIOR FILING DATE: April 10, 1997  
;; PRIOR APPLICATION NUMBER: 60/015,533  
;; PRIOR FILING DATE: April 10, 1996  
;; PRIOR APPLICATION NUMBER: 08/755,524  
;; PRIOR FILING DATE: November 22, 1996  
;; PRIOR APPLICATION NUMBER: 60/007,495  
;; PRIOR FILING DATE: November 22, 1995  
;; PRIOR APPLICATION NUMBER: 09/021,031  
;; PRIOR FILING DATE: February 10, 1998  
;; PRIOR APPLICATION NUMBER: 60/039,325  
;; PRIOR FILING DATE: February 13, 1997  
;; PRIOR APPLICATION NUMBER: 09/035,1972  
;; PRIOR FILING DATE: March 4, 1998  
;; PRIOR APPLICATION NUMBER: 60/040,431  
;; PRIOR FILING DATE: March 5, 1997  
;; PRIOR APPLICATION NUMBER: 09/041,894  
;; PRIOR FILING DATE: March 12, 1998  
;; PRIOR APPLICATION NUMBER: 60/040,199  
;; PRIOR FILING DATE: March 14, 1997  
;; PRIOR APPLICATION NUMBER: 09/050,817  
;; PRIOR FILING DATE: March 30, 1998  
;; PRIOR APPLICATION NUMBER: 60/043,792  
;; PRIOR FILING DATE: April 11, 1997  
;; PRIOR APPLICATION NUMBER: 09/074,999  
;; PRIOR FILING DATE: May 8, 1998  
;; PRIOR APPLICATION NUMBER: 60/048,431  
;; PRIOR FILING DATE: May 29, 1997  
;; PRIOR APPLICATION NUMBER: 09/107,592  
;; PRIOR FILING DATE: June 30, 1998  
;; PRIOR APPLICATION NUMBER: 60/052,751

;; PRIOR FILING DATE: July 1, 1997  
;; PRIOR APPLICATION NUMBER: 09/094,079  
;; PRIOR FILING DATE: June 9, 1998  
;; PRIOR APPLICATION NUMBER: 60/049,975  
;; PRIOR FILING DATE: June 13, 1997  
;; SOFTWARE: PERL Program  
;; SEQ ID NO 2398  
;; LENGTH: 252  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc.feature  
;; OTHER INFORMATION: Incyte ID No: hu00275763  
;; NAME/KEY: unsure  
;; LOCATION: 198, 249  
;; OTHER INFORMATION: a, t, c, g, or other  
US-09-540-2108-2398

Query Match 68.6%; Score 14.4; DB 5; Length 252;  
Best Local Similarity 93.8%; Pred. No. 1.6e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 3 tgcctgagacgctg 18  
|||||  
Db 215 tgcctgagacgctg 200

RESULT 15  
US-10-113-872-45  
;; Sequence 45, Application US/10113872  
;; GENERAL INFORMATION:  
;; APPLICANT: Watanabe, Yoshihiro  
;; APPLICANT: Henderson, Robert A.  
;; APPLICANT: Kalos, Michael D.  
;; APPLICANT: Sleath, Paul R.  
;; APPLICANT: Vedvick, Thomas S.  
;; APPLICANT: Carter, Darrick  
;; APPLICANT: Fanger, Gary R.  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
;; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
;; FILE REFERENCE: 210121.478C19  
;; CURRENT APPLICATION NUMBER: US/10/113,872  
;; CURRENT FILING DATE: 2002-03-28  
;; NUMBER OF SEQ ID NOS: 2011  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 45  
;; LENGTH: 494  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-113-872-45

Query Match 68.6%; Score 14.4; DB 6; Length 494;  
Best Local Similarity 93.8%; Pred. No. 1.7e+02;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 3 tgcctgagacgctg 18  
|||||  
Db 273 tgcctgagacgctg 288

Search completed: May 17, 2002, 17:58:45  
Job time: 9179 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 17:56:00 ; Search time 5965.48 Seconds  
(without alignments)  
76.160 Million cell updates/sec

Title: US-09-913-325-5  
Perfect score: 21  
Sequence: 1 atgtctgagaccgtctgctc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues  
Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

- 1: /cgn2\_6/ptodata/2/pna/PCRNUS.COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pna/US06.COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/pna/US07.COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/pna/US08.COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pna/US081.COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/pna/US082.COMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pna/US083.COMB.seq:\*
- 8: /cgn2\_6/ptodata/2/pna/US084.COMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pna/US085.COMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pna/US086.COMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pna/US087.COMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pna/US088.COMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pna/US089.COMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pna/US090.COMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pna/US091.COMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pna/US092.COMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pna/US093.COMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pna/US094.COMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pna/US095A.COMB.seq:\*
- 20: /cgn2\_6/ptodata/2/pna/US095B.COMB.seq:\*
- 21: /cgn2\_6/ptodata/2/pna/US095C.COMB.seq:\*
- 22: /cgn2\_6/ptodata/2/pna/US095D.COMB.seq:\*
- 23: /cgn2\_6/ptodata/2/pna/US095A.COMB.seq:\*
- 24: /cgn2\_6/ptodata/2/pna/US096A.COMB.seq:\*
- 25: /cgn2\_6/ptodata/2/pna/US096C.COMB.seq:\*
- 26: /cgn2\_6/ptodata/2/pna/US096D.COMB.seq:\*
- 27: /cgn2\_6/ptodata/2/pna/US096E.COMB.seq:\*
- 28: /cgn2\_6/ptodata/2/pna/US097A.COMB.seq:\*
- 29: /cgn2\_6/ptodata/2/pna/US097B.COMB.seq:\*
- 30: /cgn2\_6/ptodata/2/pna/US097C.COMB.seq:\*
- 31: /cgn2\_6/ptodata/2/pna/US098A.COMB.seq:\*
- 32: /cgn2\_6/ptodata/2/pna/US098B.COMB.seq:\*
- 33: /cgn2\_6/ptodata/2/pna/US098C.COMB.seq:\*
- 34: /cgn2\_6/ptodata/2/pna/US099A.COMB.seq:\*
- 35: /cgn2\_6/ptodata/2/pna/US099B.COMB.seq:\*
- 36: /cgn2\_6/ptodata/2/pna/US099C.COMB.seq:\*
- 37: /cgn2\_6/ptodata/2/pna/US100.COMB.seq:\*
- 38: /cgn2\_6/ptodata/2/pna/US101.COMB.seq:\*
- 39: /cgn2\_6/ptodata/2/pna/US6000.COMB.seq:\*
- 40: /cgn2\_6/ptodata/2/pna/US6001.COMB.seq:\*
- 41: /cgn2\_6/ptodata/2/pna/US6002.COMB.seq:\*
- 42: /cgn2\_6/ptodata/2/pna/US6003.COMB.seq:\*
- 43: /cgn2\_6/ptodata/2/pna/US6004.COMB.seq:\*

SUMMARIES				Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	
Result No.	Score	Query Match	Length DB	ID	Description
44:	100.0	21	1	PCT-US000-04875-5	Sequence 5, Appl1
45:	100.0	21	34	US-09-913-325-5	Sequence 5, Appl1
46:	100.0	21	1	US-09-944-326-5	Sequence 5, Appl1
47:	100.0	21	37	US-10-080-794-5	Sequence 5, Appl1
48:	100.0	21	100.0		Sequence 4414, Ap
49:	100.0	21	100.0		Sequence 10776, A
50:	100.0	21	100.0		Sequence 10325, A
51:	100.0	21	100.0		Sequence 10415, A
52:	100.0	21	100.0		Sequence 10412, A
53:	100.0	21	100.0		Sequence 10340, A
54:	100.0	21	100.0		Sequence 10381, A
55:	100.0	21	100.0		Sequence 10340, A
56:	100.0	21	100.0		Sequence 10340, A
57:	100.0	21	100.0		Sequence 10340, A
58:	100.0	21	100.0		Sequence 10340, A
59:	100.0	21	100.0		Sequence 10340, A
60:	100.0	21	100.0		Sequence 10340, A
61:	100.0	21	100.0		Sequence 10340, A
62:	100.0	21	100.0		Sequence 10340, A
63:	100.0	21	100.0		Sequence 10340, A
64:	100.0	21	100.0		Sequence 10340, A
65:	100.0	21	100.0		Sequence 10340, A
66:	100.0	21	100.0		Sequence 10340, A
67:	100.0	21	100.0		Sequence 10340, A
68:	100.0	21	100.0		Sequence 10340, A
69:	100.0	21	100.0		Sequence 10340, A
70:	100.0	21	100.0		Sequence 10340, A
71:	100.0	21	100.0		Sequence 10340, A
72:	100.0	21	100.0		Sequence 10340, A
73:	100.0	21	100.0		Sequence 10340, A
74:	100.0	21	100.0		Sequence 10340, A
75:	100.0	21	100.0		Sequence 10340, A

```
c 32 21 100.0 235 20 US-09-534-840-10420 Sequence 10420, A
c 33 21 100.0 235 20 US-09-534-840-10765 Sequence 10765, A
c 34 21 100.0 237 20 US-09-534-840-10384 Sequence 10384, A
c 35 21 100.0 237 20 US-09-534-840-10442 Sequence 10442, A
c 36 21 100.0 238 20 US-09-534-840-10306 Sequence 10306, A
c 37 21 100.0 240 20 US-09-534-840-10376 Sequence 10376, A
c 38 21 100.0 241 20 US-09-534-840-10351 Sequence 10351, A
c 39 21 100.0 245 20 US-09-534-840-10427 Sequence 10427, A
c 40 21 100.0 247 20 US-09-534-840-10332 Sequence 10332, A
c 41 21 100.0 248 20 US-09-534-840-10372 Sequence 10372, A
c 42 21 100.0 248 20 US-09-534-840-10375 Sequence 10375, A
c 43 21 100.0 248 20 US-09-534-840-10764 Sequence 10764, A
c 44 21 100.0 251 20 US-09-534-840-10397 Sequence 10397, A
c 45 21 100.0 253 20 US-09-534-840-9753 Sequence 9753, Ap
```

## ALIGNMENTS

```
RESULT 1
PCT-US00-04875-5
; Sequence 5, Application PC/TUS0004875
; GENERAL INFORMATION:
; APPLICANT: Gleave, Martin
; APPLICANT: Rennie, Paul S.
; APPLICANT: Miyake, Hideaki
; APPLICANT: Nelson, Colleen
; TITLE OF INVENTION: ANTISENSE THERAPY FOR CANCER
; FILE REFERENCE: UBC.P-020-WO
; CURRENT APPLICATION NUMBER: PCT/US00/04875
; CURRENT FILING DATE: 2000-02-25
; EARLIER APPLICATION NUMBER: 60/121,726
; EARLIER FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 5
; LENGTH: 21
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: antisense TRPM-2 ODN
PCT-US00-04875-5
```

```
Query Match 100.0%; Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 atgtctgagacgctctgctc 21
Db 1 atgtctgagacgctctgctc 21
```

```
RESULT 2
US-09-913-325-5
; Sequence 5, Application US/09913325
; GENERAL INFORMATION:
; APPLICANT: Gleave, Martin
; APPLICANT: Rennie, Paul S.
; APPLICANT: Miyake, Hideaki
; APPLICANT: Nelson, Colleen
; TITLE OF INVENTION: TRPM-2 ANTISENSE THERAPY
; FILE REFERENCE: UBC.P-020-WO
; CURRENT APPLICATION NUMBER: US/09/913,325
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/121,726
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 5
; LENGTH: 21
; TYPE: DNA
; ORGANISM: HUMAN
```

```
; FEATURE:
; OTHER INFORMATION: antisense TRPM-2 ODN
US-09-913-325-5
```

```
Query Match 100.0%; Score 21; DB 34; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 atgtctgagacgctctgctc 21
Db 1 atgtctgagacgctctgctc 21
```

```
RESULT 3
US-09-944-326-5
; Sequence 5, Application US/09944326
; GENERAL INFORMATION:
; APPLICANT: Gleave, Martin
; APPLICANT: Rennie, Paul S.
; APPLICANT: Miyake, Hideaki
; APPLICANT: Nelson, Colleen
; TITLE OF INVENTION: TRPM-2 ANTISENSE THERAPY
; FILE REFERENCE: UBC.P-020-2
; CURRENT APPLICATION NUMBER: US/09/944,326
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 60/121,726
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 09/913,325
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 5
; LENGTH: 21
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: antisense TRPM-2 ODN
US-09-944-326-5
```

```
Query Match 100.0%; Score 21; DB 35; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 atgtctgagacgctctgctc 21
Db 1 atgtctgagacgctctgctc 21
```

```
RESULT 4
US-10-080-794-5
; Sequence 5, Application US/10080794
; GENERAL INFORMATION:
; APPLICANT: Gleave, Martin
; APPLICANT: Rennie, Paul S.
; APPLICANT: Miyake, Hideaki
; APPLICANT: Nelson, Colleen
; APPLICANT: Monia, Brett P.
; TITLE OF INVENTION: TRPM-2 ANTISENSE THERAPY USING AN OLIGONUCLEOTIDE
; FILE REFERENCE: UBC.P-020-3
; CURRENT APPLICATION NUMBER: US/10/080,794
; CURRENT FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: 60/121,726
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 09/913,325
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 09/944,326
; PRIOR FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 5
```



```
LENGTH: 21
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: anltsense TRPM-2 ODN
US-10-080-794-5
```

```
Query Match          100.0%; Score 21; DB 37; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 attgtctgagaccgtctgctc 21
    |||||
Db 1 attgtctgagaccgtctgctc 21
```

```
RESULT 5
US-09-716-472-4414/c
; Sequence 4414, Application US/09716472
; GENERAL INFORMATION:
; APPLICANT: Sllas-Santiago, Immaculada
; APPLICANT: Distefano, Peter
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2033-001
; CURRENT APPLICATION NUMBER: US/09/716,472
; CURRENT FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/166,601
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 6233
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4414
; LENGTH: 168
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(168)
; OTHER INFORMATION: n = A,T,C or G
US-09-716-472-4414
```

```
Query Match          100.0%; Score 21; DB 28; Length 168;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 attgtctgagaccgtctgctc 21
    |||||
Db 112 ATTGCTGAGACCGCTCTGCTC 92
```

```
RESULT 6
US-09-534-840-10776/c
; Sequence 10776, Application US/09534840
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MOLE
; FILE REFERENCE: PD-1010 CIP
; CURRENT APPLICATION NUMBER: US/09/534,840
; CURRENT FILING DATE: 1992-07-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 13675
; SOFTWARE: PERL Program
; SEQ ID NO 10776
; LENGTH: 173
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu01344983
US-09-534-840-10776
```

```
Query Match          100.0%; Score 21; DB 20; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 attgtctgagaccgtctgctc 21
    |||||
Db 77 ATTGCTGAGACCGCTCTGCTC 57
```

```
RESULT 7
US-09-534-840-10325/c
; Sequence 10325, Application US/09534840
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MO
; FILE REFERENCE: PD-1010 CIP
; CURRENT APPLICATION NUMBER: US/09/534,840
; CURRENT FILING DATE: 1992-07-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 13675
; SOFTWARE: PERL Program
; SEQ ID NO 10325
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu01113942
US-09-534-840-10325
```

```
Query Match          100.0%; Score 21; DB 20; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 attgtctgagaccgtctgctc 21
    |||||
Db 154 ATTGCTGAGACCGCTCTGCTC 134
```

```
RESULT 8
US-09-534-840-10415/c
; Sequence 10415, Application US/09534840
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MO
; FILE REFERENCE: PD-1010 CIP
; CURRENT APPLICATION NUMBER: US/09/534,840
; CURRENT FILING DATE: 1992-07-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 13675
; SOFTWARE: PERL Program
; SEQ ID NO 10415
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

NAME/KEY: misc-feature  
OTHER INFORMATION: Incyte ID No: hu01055183  
NAME/KEY: unsure  
LOCATION: 17, 40  
OTHER INFORMATION: a, t, c, g, or other  
US-09-534-840-10415

Query Match 100.0%; Score 21; DB 20; Length 177;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagaccgtctgtc 21  
|||||  
Db 154 ATTGCTGAGACCGTCTGTGTC 134

## RESULT 9

US-09-534-840-10784/c  
Sequence 10784, Application US/09534840  
GENERAL INFORMATION:  
APPLICANT: Sellhammer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullaly, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MOLEC  
FILE REFERENCE: PD-1010 CIP  
CURRENT APPLICATION NUMBER: US/09/534,840  
CURRENT FILING DATE: 1992-07-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 13675  
SOFTWARE: PERL Program  
SEQ ID NO 10784  
LENGTH: 190  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc-feature  
OTHER INFORMATION: Incyte ID No: hu01342214  
US-09-534-840-10784

Query Match 100.0%; Score 21; DB 20; Length 190;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagaccgtctgtc 21  
|||||  
Db 154 ATTGCTGAGACCGTCTGTGTC 134

## RESULT 10

US-09-534-840-10412/c  
Sequence 10412, Application US/09534840  
GENERAL INFORMATION:  
APPLICANT: Sellhammer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullaly, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MOLEC  
FILE REFERENCE: PD-1010 CIP  
CURRENT APPLICATION NUMBER: US/09/534,840  
CURRENT FILING DATE: 1992-07-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 13675  
SOFTWARE: PERL Program  
SEQ ID NO 10412  
LENGTH: 192  
TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc-feature  
OTHER INFORMATION: Incyte ID No: hu01085063  
NAME/KEY: unsure  
LOCATION: 119, 124, 134, 175  
OTHER INFORMATION: a, t, c, g, or other  
US-09-534-840-10412

Query Match 100.0%; Score 21; DB 20; Length 192;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagaccgtctgtc 21  
|||||  
Db 76 ATTGCTGAGACCGTCTGTGTC 56

## RESULT 11

US-08-485-657-14/c  
Sequence 14, Application US/08485657  
GENERAL INFORMATION:  
APPLICANT: Kirschling, Deborah J  
APPLICANT: Gudkov, Andrei  
APPLICANT: Roninson, Igor B  
TITLE OF INVENTION: Genes and Genetic Elements Associated  
TITLE OF INVENTION: with Sensitivity to Cisplatin  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESS: ALLEGRETTI & WITCOFF, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,657  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/199,900  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Noonan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 93,354-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 910-221-5317  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 195 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-485-657-14

Query Match 100.0%; Score 21; DB 8; Length 195;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 attgtctgagaccgtctgtc 21  
|||||  
Db 153 ATTGCTGAGACCGTCTGTGTC 133

```
RESULT 12
US-09-366-380-14/C
; Sequence 14, Application US/09366380
; GENERAL INFORMATION:
; APPLICANT: Kirschling, Deborah J
; APPLICANT: Gudkov, Andrei
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Genes and Genetic Elements Associated
; TITLE OF INVENTION: With Sensitivity to Cisplatin
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/366,380
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/199,900
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noonan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-N
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-366-380-14

Query Match          100.0%; Score 21; DB 17; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtctgaagacgctcgtc 21
   ||||||||||||||||
DB 153 ATTGCTGAGACCGCTGCTC 133

RESULT 13
US-09-534-840-9752/C
; Sequence 9752, Application US/09534840
; GENERAL INFORMATION:
; APPLICANT: Sellhammer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MOED
; FILE REFERENCE: PD-1010 CIP
; CURRENT APPLICATION NUMBER: US/09/534,840
; CURRENT FILING DATE: 1992-07-17
; Prior application data removed - consult PALM or file wrapper
```

```
NUMBER OF SEQ ID NOS: 13675
; SOFTWARE: PERL Program
; SEQ ID NO 9752
; LENGTH: 199
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu01227727
; NAME/KEY: unsure
; LOCATION: 118
; OTHER INFORMATION: a, t, c, g, or other
US-09-534-840-9752

Query Match          100.0%; Score 21; DB 20; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtctgaagacgctcgtc 21
   ||||||||||||||||
DB 153 ATTGCTGAGACCGCTGCTC 133

RESULT 14
US-09-534-840-10336/C
; Sequence 10336, Application US/09534840
; GENERAL INFORMATION:
; APPLICANT: Sellhammer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MO
; FILE REFERENCE: PD-1010 CIP
; CURRENT APPLICATION NUMBER: US/09/534,840
; CURRENT FILING DATE: 1992-07-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 13675
; SOFTWARE: PERL Program
; SEQ ID NO 10336
; LENGTH: 203
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu01308898
US-09-534-840-10336

Query Match          100.0%; Score 21; DB 20; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtctgaagacgctcgtc 21
   ||||||||||||||||
DB 154 ATTGCTGAGACCGCTGCTC 134

RESULT 15
US-09-534-840-9756/C
; Sequence 9756, Application US/09534840
; GENERAL INFORMATION:
; APPLICANT: Sellhammer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MO
; FILE REFERENCE: PD-1010 CIP
; CURRENT APPLICATION NUMBER: US/09/534,840
```

; CURRENT FILING DATE: 1992-07-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 13675  
; SOFTWARE: PERL Program  
; SEQ ID NO 9756  
; LENGTH: 205  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: hu01099825  
US-09-534-840-9756

Query Match 100.0%; Score 21; DB 20; Length 205;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 attgtctgagacgctcgtc 21  
|||||  
Db 136 ATGCTGTGAGACCGCTGTGTC 116

Search completed: May 17, 2002, 17:56:00  
Job time: 12667 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:06:10 ; Search time 111.1 Seconds  
(Without alignments)  
46.429 Million cell updates/sec

Title: US-09-913-325-5

Sequence: 1 atgtctgagacgctctgtc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0 .

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/pdata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/pdata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/pdata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/pdata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/pdata/2/ina/PCRTUS.COMB.seq:\*  
6: /cgn2\_6/pdata/2/ina/Backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	21	100.0	195	2	US-08-485-657A-14
C 2	21	100.0	195	5	PCR-US95-02303-14
C 3	19.4	92.4	181	2	US-08-485-657A-19
C 4	19.4	92.4	181	5	PCR-US95-02303-18
C 5	15.2	72.4	6671	1	US-08-280-443-1
C 6	15.2	72.4	6671	1	US-08-457-459-1
C 7	15.2	72.4	6671	1	US-08-555-678-1
C 8	15.2	72.4	6671	5	PCR-US95-02275-1
C 9	15.2	72.4	6671	1	US-08-151-574-31
C 10	15.2	72.4	6756	2	US-08-419-448-31
C 11	15.2	72.4	6756	4	US-09-233-510-31
C 12	14.8	70.5	1438	3	US-09-187-331-4
C 13	14.8	70.5	1438	4	US-09-470-946-4
C 14	14.8	70.5	3563	4	US-09-041-886-20
C 15	14.8	70.5	3596	2	US-08-779-801-5
C 16	14.8	70.5	3596	4	US-09-298-441-5
C 17	14.6	69.5	2820	5	PCR-US93-11725-1
C 18	14.6	69.5	2822	1	US-08-015-985-4
C 19	14.4	68.6	2002	2	US-08-715-202A-3
C 20	14.2	67.6	331	4	US-09-060-756-310
C 21	14.2	67.6	939	4	US-09-105-390-45
C 22	14.2	67.6	944	2	US-08-522-421-2
C 23	14.2	67.6	1020	4	US-09-105-390-61
C 24	14.2	67.6	1130	1	US-07-864-004B-1
C 25	14.2	67.6	1130	1	US-08-251-937A-1
C 26	14.2	67.6	1130	1	US-08-212-133A-5
C 27	14.2	67.6	1130	1	US-08-474-503-3

C 28	14.2	67.6	1130	2	US-08-670-707A-3	Sequence 3, Appl
C 29	14.2	67.6	1130	4	US-09-037-601-3	Sequence 3, Appl
C 30	14.2	67.6	1130	5	PCR-US93-03275-1	Sequence 1, Appl
C 31	14.2	67.6	1130	5	PCR-US94-13200-3	Sequence 3, Appl
C 32	14.2	67.6	1428	2	US-08-786-999-2	Sequence 2, Appl
C 33	14.2	67.6	1428	4	US-08-969-987-8	Sequence 8, Appl
C 34	14.2	67.6	1494	1	US-08-585-758A-3	Sequence 3, Appl
C 35	14.2	67.6	1494	1	US-08-977-818-3	Sequence 3, Appl
C 36	14.2	67.6	1494	2	US-08-670-274B-3	Sequence 3, Appl
C 37	14.2	67.6	1494	1	US-09-146-187-3	Sequence 3, Appl
C 38	14.2	67.6	1623	1	US-08-121-202-3	Sequence 3, Appl
C 39	14.2	67.6	2200	1	US-08-592-126-102	Sequence 102, App
C 40	14.2	67.6	2512	4	US-09-105-390-7	Sequence 7, Appl
C 41	14.2	67.6	4334	2	US-08-670-707A-38	Sequence 38, Appl
C 42	14.2	67.6	4334	4	US-09-037-601-38	Sequence 38, Appl
C 43	14.2	67.6	6402	2	US-08-670-707A-36	Sequence 36, Appl
C 44	14.2	67.6	6402	4	US-09-037-601-36	Sequence 36, Appl
C 45	14.2	67.6	7032	4	US-09-324-867-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-08-485-657A-14/c  
Sequence 14, Application US/08485657A  
Patent No. 5942389  
GENERAL INFORMATION:  
APPLICANT: Kirschling, Deborah J  
APPLICANT: Gudkov, Andrei  
APPLICANT: Roninson, Igor B  
TITLE OF INVENTION: Genes and Genetic Elements Associated  
TITLE OF INVENTION: With Sensitivity to Cisplatin  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
STREET: 300 South Wacker Drive, 32nd Floor  
City: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,657A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5942389nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 93,354-N  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
TELEX:  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 195 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-485-657A-14  
Query Match 100.0%; Score 21; DB 2; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.027;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 153 ATGTCTGAGACCGTCTGTC 133

## RESULT 2

PCT-US95-02303-14/c  
; Sequence 14, Application PC/TUS9502303  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Genes and Genetic Elements Associated  
; TITLE OF INVENTION: With Sensitivity to Cisplatin  
; NUMBER OF SEQUENCES: 25  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02303  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 195 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
PCT-US95-02303-14

Query Match 100.0%; Score 21; DB 5; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.027;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 attgtctgagaccgtctgtc 21  
|||  
Db 153 ATGTCTGAGACCGTCTGTC 133

## RESULT 3

US-08-485-657A-19/c  
; Sequence 19, Application US/08485657A  
; Patent No. 5942389  
; GENERAL INFORMATION:  
; APPLICANT: Kirschling, Deborah J  
; APPLICANT: Gudkov, Andrei  
; APPLICANT: Roninson, Igor B  
; TITLE OF INVENTION: Genes and Genetic Elements Associated  
; TITLE OF INVENTION: With Sensitivity to Cisplatin  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,657A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: B00  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5942389nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 93,354-N  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; TELEX:

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 181 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
US-08-485-657A-19

Query Match 92.4%; Score 19.4; DB 2; Length 181;  
Best Local Similarity 95.2%; Pred. No. 0.18;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 attgtctgagaccgtctgtc 21  
|||  
Db 153 ATGTCTGAGACCGTCTGTC 119

## RESULT 4

PCT-US95-02303-18/c  
; Sequence 18, Application PC/TUS9502303  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Genes and Genetic Elements Associated  
; TITLE OF INVENTION: With Sensitivity to Cisplatin  
; NUMBER OF SEQUENCES: 25  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02303  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 181 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
PCT-US95-02303-18

Query Match 92.4%; Score 19.4; DB 5; Length 181;  
Best Local Similarity 95.2%; Pred. No. 0.18;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 attgtctgagaccgtctgtc 21  
|||  
Db 153 ATGTCTGAGACCGTCTGTC 119

## RESULT 5

US-08-280-443-1/c  
; Sequence 1, Application US/08280443  
; Patent No. 5643778  
; GENERAL INFORMATION:  
; APPLICANT: Nishikura, Kazuko  
; TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use  
; TITLE OF INVENTION: Thereof  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/280,443
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/197,794
FILING DATE: 17-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST49AUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6671 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 155..3832
US-08-280-443-1

Query Match      72.4% Score 15.2; DB 1; Length 6671;
Best Local Similarity 85.0%; Pred. No. 52;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      2 ttgtctgagaccgtctgtc 21
      ||| || ||||| ||||| |||
DB      808 ttgcgtatgaccgtctgtc 789

RESULT      6
US-08-457-459-1/c
Sequence 1, Application US/08457459
Patent No. 5677428
GENERAL INFORMATION:
APPLICANT: Nishikura, Kazuko
TITLE OF INVENTION: RNA Editing Enzyme and Methods of Use
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P.O. Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,459
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/197,794
FILING DATE: 17-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/280,443
FILING DATE: 25-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST49CUSA
TELECOMMUNICATION INFORMATION:
```

```
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6671 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 155..3832
US-08-457-459-1

Query Match      72.4% Score 15.2; DB 1; Length 6671;
Best Local Similarity 85.0%; Pred. No. 52;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      2 ttgtctgagaccgtctgtc 21
      ||| || ||||| ||||| |||
DB      808 ttgcgtatgaccgtctgtc 789

RESULT      7
US-08-555-678-1/c
Sequence 1, Application US/08555678
Patent No. 5763174
GENERAL INFORMATION:
APPLICANT: Nishikura, Kazuko
TITLE OF INVENTION: RNA Editing Enzyme and Methods
TITLE OF INVENTION: of Use Thereof
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, P.O. Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/555,678
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/197,794
FILING DATE: 17-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/280,443
FILING DATE: 25-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,459
FILING DATE: 01-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST49DUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9206
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6671 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
```







```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6756 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Aspergillus ficuum (Aspergillus niger)
; STRAIN: NRRL 3135
; IMMEDIATE SOURCE:
; LIBRARY: Lambda AF
; CLONE: pAF2-3, pAF2-6, pAF2-7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 210..253
; FEATURE:
; NAME/KEY: Intron
; LOCATION: 254..355
; FEATURE:
; NAME/KEY: exon
; LOCATION: 356..1715
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(210..253, 356..1715)
; OTHER INFORMATION: /codon_start= 210
; OTHER INFORMATION: /product= "Phytase"
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 210..380
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 381..1712
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "inositol phosphate
; OTHER INFORMATION: /phosphatase"
; OTHER INFORMATION: /product= "Phytase"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
;
US-09-233-510-31

Query Match 72.4%; Score 15.2; DB 4; Length 6756;
Best Local Similarity 85.0%; Pred. No. 52;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ttgtctgagaccgtctgtc 21
|||1111111111111111
DB 130 TTATATGAGACCGTCGGGTC 149

RESULT 12
US-09-187-331-4/c
; Sequence 4, Application US/09187331
; Patent No. 6043056
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Gorgone, Gina A.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
; FILE REFERENCE: PF-0631 US
; CURRENT APPLICATION NUMBER: US/09/187,331
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1438
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2705267
;
US-09-187-331-4

Query Match 70.5%; Score 14.8; DB 3; Length 1438;
Best Local Similarity 88.9%; Pred. No. 65;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 gctctgagaccgtctgtc 21
|||||1111111111111111
DB 1232 GTCGTGAGACATCTGTC 1215

RESULT 13
US-09-470-946-4/c
; Sequence 4, Application US/09470946
; Patent No. 6358923
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Gorgone, Gina A.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
; FILE REFERENCE: PF-0631 US
; CURRENT APPLICATION NUMBER: US/09/470,946
; CURRENT FILING DATE: 1999-12-22
; EARLIER APPLICATION NUMBER: US 09/187,331
; EARLIER FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1438
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2705267
;
US-09-470-946-4

Query Match 70.5%; Score 14.8; DB 4; Length 1438;
Best Local Similarity 88.9%; Pred. No. 65;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 gctctgagaccgtctgtc 21
|||||1111111111111111
DB 1232 GTCGTGAGACATCTGTC 1215

RESULT 14
US-09-041-886-20/c
; Sequence 20, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabinzaden, Sharoz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-1J 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 3563 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 3..3550
US-09-041-886-20

Query Match          70.5%; Score 14.8; DB 4; Length 3563;
Best Local Similarity 88.9%; Pred. No. 76;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 tgctgagacgctcgt 20
   |||||  |||||
Db 2385 TGCTGAGATGCTCTGCT 2368

RESULT 15
US-08-779-801-5/c
Sequence 5, Application US/08779801
Patent No. 5853995
GENERAL INFORMATION:
APPLICANT: Lee, Cheng-Chi
TITLE OF INVENTION: Large Scale Genotyping of
TITLE OF INVENTION: Diseases and a Diagnostic Test for Spinocerebellar Ataxia Type
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Benjamin Aaron Adler, Ph.D., J.D.
STREET: 8011 Candle Lane
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word for Macintosh
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/779,801
FILING DATE: January 7, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benjamin Aaron Adler, Ph.D., J.D.
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5968
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3596
TYPE: nucleic acid

```

```

STRANDEDNESS: double-stranded
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: human
TISSUE TYPE: brain
IMMEDIATE SOURCE:
LIBRARY: Primary human brain cDNA
CLONE: BI-1(V2)-GGCAG
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 19p13
US-08-779-801-5

Query Match          70.5%; Score 14.8; DB 2; Length 3596;
Best Local Similarity 88.9%; Pred. No. 76;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 tgctgagacgctcgt 20
   |||||  |||||
Db 2385 TGCTGAGATGCTCTGCT 2368

Search completed: May 17, 2002, 16:06:12
Job time: 12746 sec

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:13:54 ; Search time 447.74 Seconds  
(without alignments)  
80.527 Million cell updates/sec

Title: US-09-913-325-5

Sequence: 1 attgtctgagaccgtctgttc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq\_032802.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	21	AAA94227	Human testostero
2	21	100.0	195	AAT00416	Genetic suppressor
3	21	100.0	396	AAF94824	Human ovarian can
4	21	100.0	491	AAO3751	Human secreted pro
5	21	100.0	512	AAV89150	EST Clone CGI. Ho
6	21	100.0	572	AAZ42136	Human normal blad
7	21	100.0	922	AAA43857	Human secreted exp
8	21	100.0	1568	AA544948	cdna encoding nove
9	21	100.0	1651	AAO11503	Cytolysis Inhibito

C 10	21	100.0	1678	22	AAH23086	Osteoarthritis tis
C 11	17.8	84.8	518	22	AAD05495	Human secreted pro
C 12	17.8	84.8	518	22	AAD05522	Human secreted pro
C 13	17.8	84.8	1670	22	AAD08210	Human secreted pro
C 14	17.8	84.8	1670	22	AAD05523	Human secreted pro
C 15	17.8	84.8	1684	22	AAD08191	Human secreted pro
C 16	16.2	77.1	2304	22	AA545136	cdna encoding nove
C 17	16.2	77.1	325791	22	AA543104	Human Oestrogen re
C 18	16	76.2	11674	22	AAK70965	Human Immune/haema
C 19	15.8	75.2	355	22	AAH34666	Human colon cancer
C 20	15.8	75.2	1580	22	AA530562	DNA encoding novel
C 21	15.8	75.2	1580	22	AAO06615	Human reproductive
C 22	15.8	75.2	8888	20	AAH02665	BP-692047, Seq ID 4
C 23	15.4	73.3	374	22	AA532978	DNA encoding CARDI
C 24	15.4	73.3	2097	23	ABL05375	Drosophila melanog
C 25	15.4	73.3	4284	23	ABL05374	Drosophila melanog
C 26	15.4	73.3	6245	23	ABL19688	Drosophila melanog
C 27	15.4	73.3	6632	23	ABL19686	Drosophila melanog
C 28	15.4	73.3	6655	23	ABL05612	Drosophila melanog
C 29	15.2	72.4	51	22	AAH79918	Human DNA containi
C 30	15.2	72.4	332	22	AAO1942	Human reproductive
C 31	15.2	72.4	342	22	ABA71312	Human foetal liver
C 32	15.2	72.4	342	22	AAK19616	Human brain expres
C 33	15.2	72.4	342	22	AAK45622	Human bone marrow
C 34	15.2	72.4	342	22	AAI51555	Probe #20241 used
C 35	15.2	72.4	366	21	AAAC48662	Arabidopsis thalia
C 36	15.2	72.4	367	21	AAAC34794	Human foetal liver
C 37	15.2	72.4	484	22	AA558577	Human brain expres
C 38	15.2	72.4	484	22	AAK06699	Human bone marrow
C 39	15.2	72.4	484	22	AAK32397	Probe #6926 used t
C 40	15.2	72.4	484	22	AAI38240	DNA encoding novel
C 41	15.2	72.4	579	23	AA565858	DNA encoding novel
C 42	15.2	72.4	1466	23	AA580037	DNA encoding novel
C 43	15.2	72.4	1668	23	AA592996	DNA encoding novel
C 44	15.2	72.4	1668	23	AA593802	DNA encoding novel
C 45	15.2	72.4	1803	22	AAO05175	Human secreted pro

## ALIGNMENTS

AAA94227	1	standard; DNA; 21 BP.
AAA94227	1	12-JAN-2001 (first entry)
AAA94227	1	Human testostero
AAA94227	1	Human: testostero
AAA94227	1	sulfated glycoprotein-2; SGP-2; cancer; antisense oligonucleotide; ss.
AAA94227	1	Homo sapiens.
AAA94227	1	WO200049937-A2.
AAA94227	1	31-AUG-2000.
AAA94227	1	25-FEB-2000; 2000WO-US04875.
AAA94227	1	26-FEB-1999; 99US-0121726.
AAA94227	1	(UYBR-) UNIV BRITISH COLUMBIA.
AAA94227	1	Gleaves M, Rennie PS, Miyake H, Nelson C;
AAA94227	1	WPI: 2000-533132/48.
AAA94227	1	Treating prostatic tumors and renal cancers by antisense inhibition of
AAA94227	1	the testostero
AAA94227	1	repressed prostate messenger-2 gene -

PS Claim 4; Page 36; 38pp; English.

XX The present sequence is an antisense oligonucleotide directed at the

CC human testosterone-repressed prostate message-2 (TRPM-2, also known as

CC clusterin, sklated glycoprotein-2 or SGP-2). The sequence was shown to

CC promote the regression of tumours, and oligonucleotides directed

CC at human TRPM-2 can be used in the treatment of tumour cells expressing

CC the TRPM-2 gene. These include prostate cancer, renal cell cancer and

CC some breast cancer cells. In addition to this, they also increase the

CC chemosensitivity of the cells, meaning that conventional chemotherapy is

CC more effective.

XX

SQ Sequence 21 BP; 3 A; 5 C; 6 G; 7 T; 0 other:

Query Match 100.0%; Score 21; DB 21; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.16;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 attgtctgagaccgtctgctgc 21

DB 1 attgtctgagaccgtctgctgc 21

RESULT 2

AAAT00416/C

ID AAT00416 standard; CDNA; 195 BP.

XX

AC AAT00416;

XX

DT 26-MAR-1996 (first entry)

XX

DE Genetic suppressor element HL7.1.

XX

KM Genetic suppressor element; GSE; platinum-based drug; cisplatin;

KM chemotherapy; HL7.1; testosterone-repressed prostatic message-2; TRPM-2;

KM ss.

XX

OS Synthetic.

XX

PN WO9522612-A2.

XX

PD 24-AUG-1995.

XX

PF 22-FEB-1995; 95WO-US02303.

XX

PR 22-FEB-1994; 94US-0199900.

XX

PA (UNIT I) UNIT ILLINOIS FOUND.

XX

PI Gudkov A, Kirschling DJ, Roninson IB;

XX

DR WPI: 1995-302718/39.

XX

PT Genetic suppressor elements which confer resistance to

PT platinum-based drugs, eg. cisplatin, on cancer cells - useful for

PT enhancement of chemotherapy, and for diagnosis of resistance to

PT these drugs.

XX

PS Claim 14; Fig 17; 75pp; English.

XX

CC The sequences represented by AAT00405-T00418 are genetic suppressor

CC elements (GSEs). This sequence represents GSE HL6.10. This sequence

CC shows homology to the CDNA encoding testosterone-repressed prostatic

CC message-2 (TRPM-2). These sequences were obtained from a CDNA library

CC derived from the total CDNA of a cisplatin sensitive cell. Genetic

CC suppressor elements confer resistance to platinum-based drugs (PDS),

CC such as cisplatin. These functional GSEs can then be used to create

CC probes for the parent gene. The probes can then be used in a method of

CC measuring the level of GSE gene expression. The GSEs can be used in

CC methods of diagnosis of resistance to PDS by measuring the level of

CC expression of GSE genes. The GSEs are also used in methods to overcome

CC resistance to PDS in cancer cells. The GSEs (or fragments of them) can

CC be used to inhibit the function of genes associated with sensitivity to

CC PDS. For enhancement of chemotherapy, a GSE can be transferred (either

CC alone or with another gene) on an expression vector into blood

CC progenitor cells from a cancer patient. The cells are returned to the

CC patients circulation and allowed to repopulate the blood before

CC aggressive chemotherapy is carried out (using higher cisplatin

CC concentrations than normal), this will thereby avoid toxic side effects

CC to the immune system as the blood cells will be GSE resistant.

XX

SQ Sequence 195 BP; 46 A; 45 C; 71 G; 33 T; 0 other:

Query Match 100.0%; Score 21; DB 16; Length 195;

Best Local Similarity 100.0%; Pred. No. 0.23;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 attgtctgagaccgtctgctgc 21

DB 153 ATTGCTGTGAGACCGTCTGCTC 133

RESULT 3

AAF94824/C

ID AAF94824 standard; CDNA; 396 BP.

XX

AC AAF94824;

XX

DT 23-MAY-2001 (first entry)

XX

DE Human ovarian cancer associated coding sequence SEQ ID NO: 15.

XX

KM Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.

XX

OS Homo sapiens.

XX

PN WO200118046-A2.

XX

PD 15-MAR-2001.

XX

PF 08-SEP-2000; 2000WO-US24827.

XX

PR 10-SEP-1999; 99US-0394374.

PR 01-MAY-2000; 2000US-0561778.

PR 15-AUG-2000; 2000US-0640173.

PR 07-SEP-2000; 2000US-0656668.

XX

PA (CORI-) CORIXA CORP.

XX

PI Xu J, Stolk JA;

XX

DR WPI: 2001-211395/21.

XX

PT Isolated polypeptides associated with ovarian carcinomas, and the

PT nucleic acids that encode them, useful for the prevention diagnosis and

PT treatment of ovarian cancers -

XX

PS Claim 18; Page 120; 189pp; English.

XX

CC The present invention provides a number of coding sequences and proteins,

CC the over-expression of which is associated with ovarian carcinoma/cancer.

CC These can be used in the diagnosis, treatment and prevention of ovarian

CC cancer, optionally by gene therapy or in the form of a vaccine. The

CC present sequence is an example of one of these sequences.

XX

SQ Sequence 396 BP; 129 A; 83 C; 121 G; 62 T; 1 other:

Query Match 100.0%; Score 21; DB 22; Length 396;

Best Local Similarity 100.0%; Pred. No. 0.26;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 attgtctgagaccgtctgctgc 21

|||||

Db 145 ATTGCTGAGACCGTCTGCTC 125

# RESULT 4

AA03751/c  
ID AAC03751 standard; cDNA: 491 BP.

XX AAC03751;

XX 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 3749.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX Homo sapiens.

XX EPI033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GENSET ) GENSET.

XX Dumas Mline Edwards J, Duclert A, Giordano J;

DR WPI: 2000-500381/45.

XX P-PSDB: AAC03745.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
diagnosing CDNs and genomic DNAs that correspond to 5' ESTs and for  
diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX Claim 1: SEQ ID 3749; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. An ORF has been identified within the  
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
CC derived from 30 different tissues. EST sequences usually correspond  
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
CC well suited for isolating cDNA sequences derived from the 5' ends of  
CC mRNAs and even in those cases where longer cDNA sequences have been  
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
CC gene therapy and chromosome mapping procedures. They are used to obtain  
CC upstream regulatory sequences and to design expression and secretion  
CC vectors.

XX Sequence 491 BP; 149 A; 114 C; 149 G; 78 T; 1 other;

Query Match 100.0%; Score 21; DB 21; Length 491;

Best Local Similarity 100.0%; Pred. No. 0.27;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtctgagaccgtctgctc 21

DB 194 ATTGCTGAGACCGTCTGCTC 174

# RESULT 5

AAV89150/c  
ID AAV89150 standard; cDNA: 512 BP.

XX AAV89150;

DT 15-FEB-1999 (first entry)

XX EST clone CGL.

XX Human; secreted protein; expressed sequence tag; EST; haematopoiesis;  
KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;  
KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;  
KW gene therapy; ss.

XX Homo sapiens.

XX W09845436-A2.

XX 15-OCT-1998.

XX 10-APR-1998; 98WO-US06955.

XX 10-APR-1997; 97US-0838821.

XX (GEMV ) GENETICS INST INC.

XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

XX Racie LA, Spaulding V, Treacy M;

XX WPI: 1999-070077/06.

PT New polynucleotides encoding human secreted proteins - derived from  
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
PT ovary, pituitary, retina and colon cDNA libraries.  
XX Claim 1: Page 126; 618pp; English.

XX The present sequence represents a human expressed sequence tag (EST).  
CC The polynucleotide, which is a secreted EST, and the encoded protein  
CC are predicted to have useful biological activities which would make  
CC them suitable for treating, preventing or ameliorating medical  
CC conditions in humans and animals, although no supporting data is  
CC given. Suggested activities include nutritional activity, immune  
CC stimulating or suppressing activity, haematopoiesis regulating  
CC activity, tissue growth activity, activin/inhibin activity,  
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic  
CC activity, receptor/ligand activity, anti-inflammatory activity,  
CC cadherin/tumour invasion suppressor activity, tumour inhibition  
CC activity. The polynucleotide may also be useful for gene therapy.

XX Sequence 512 BP; 154 A; 114 C; 150 G; 94 T; 0 other;

Query Match 100.0%; Score 21; DB 20; Length 512;

Best Local Similarity 100.0%; Pred. No. 0.27; Mismatches 0; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtctgagaccgtctgctc 21

DB 121 ATTGCTGAGACCGTCTGCTC 101

# RESULT 6

AA42136/c  
ID AA42136 standard; cDNA: 572 BP.

XX AA42136;

XX 31-JAN-2000 (first entry)

XX Human normal bladder tissue cDNA derived EST 15.

XX Human; bladder; treatment; EST; expressed sequence tag; cytostatic;

XX cancer; gene therapy; ss.

XX Homo sapiens.

XX DE19818620-A1.

PD 28-OCT-1999.  
 XX 21-APR-1998; 98DE-1018620.  
 PF 21-APR-1998; 98DE-1018620.  
 PR 21-APR-1998; 98DE-1018620.  
 XX  
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 XX WPI: 1999-602416/52.  
 DR  
 XX New polypeptides and their nucleic acids, useful for treatment of  
 PT bladder tumour and identification of therapeutic agents -  
 XX  
 PS Claim 3; Page 158; 366pp; German.  
 XX  
 CC This invention describes novel polypeptide fragment sequences (I) and  
 CC their encoding nucleic acids (II) which are highly expressed in normal  
 CC bladder tissue and have cytostatic activity. (II) are used for  
 CC recombinant expression of (I) and to isolate complete genes. (I) are  
 CC used to identify agents suitable for the treatment of bladder tumours,  
 CC to directly treat this form of cancer (including expression from gene  
 CC therapy vectors), or are used in a preparation for cancer treatment. (I)  
 CC is also used for the generation of specific antibodies. (II) are  
 CC identified by assembling ESTs (expressed sequence tags) from a  
 CC particular tissue type before comparison of expression patterns. This  
 CC allows a significantly longer fragment of the gene to be revealed, and  
 CC therefore reduces the number of failures because of ESTs from different  
 CC libraries representing different parts of the same unknown gene  
 CC distorting the estimated frequency of occurrence in a particular tissue.  
 CC AA42122-242248 represent EST fragments derived from a human normal  
 CC bladder tissue cDNA library which encode the protein fragments  
 CC represented in AAY60329-Y60591.  
 CC  
 XX Sequence 572 BP; 161 A; 128 C; 186 G; 97 T; 0 other:  
 SO

Query Match 100.0%; Score 21; DB 20; Length 572;  
 Best Local Similarity 100.0%; Pred. No. 0.28;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 attgtctgagacgctctgctc 21  
 ||||||||||||||||||||  
 DB 262 ATTGCTGTGAGACCTCTGTGTC 242

RESULT 7  
 ID AAA43857 standard; cDNA; 922 BP.  
 XX  
 AC AAA43857;  
 XX  
 DT 21-AUG-2000 (first entry)  
 XX  
 DE Human secreted expressed sequence tag SEQ ID NO:432.  
 XX  
 KW Human: mouse; chicken; rat; secreted expressed sequence tag; SESTR;  
 KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
 KW immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;  
 KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antitumour;  
 KW antiviral; antidiabetic; antischistosomal; antiparkinsonian;  
 KW anticancer; osteoprotective; neurotrophic; antiproliferative;  
 KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
 KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
 KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
 KW lymphoid dependent diabetes; burn; osteoporosis; osteoarthritis;  
 KW central nervous system disorder; Alzheimer's disease; stroke;  
 KW Parkinson's disease; Huntington's disease; coagulation disorder;  
 KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
 KW tumour; infection; depression; psoriasis; ss.  
 KW  
 XX Homo sapiens.  
 OS

XX WO200021991-A1.  
 PN  
 XX 20-APR-2000.  
 PD  
 XX 15-OCT-1999; 99WO-US24206.  
 PF 15-OCT-1999; 98US-0104436.  
 XX  
 PR 15-OCT-1999; 98US-0104436.  
 XX  
 PA (GENY) GENETICS INST INC.  
 PI Jacobs K, McCoy JM, LaValle ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Bowman MR;  
 XX WPI: 2000-317938/27.  
 DR  
 XX Isolated polynucleotides, and encoded proteins, comprising secreted  
 PT expressed sequence tags (SESTs), useful for treating various disorders  
 PT such as autoimmune, infectious, and central nervous system disorders -  
 XX  
 PS Claim 1; Page 316; 803pp; English.  
 XX  
 CC AAA43426 to AAA45925 represent specifically claimed secreted expressed  
 CC sequence tags (SESTs), isolated from human, mouse, chicken and rat  
 CC tissue sources. The SESTs can have a range of activities depending on  
 CC the tissues they were isolated from. The activities include:  
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
 CC chemokine; analgesic; haemostatic; thrombolytic; antiinflammatory;  
 CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;  
 CC antischistosomal; antiparkinsonian; antiproliferative;  
 CC neurotrophic; antiparkinsonian; antiproliferative; cerebroprotective;  
 CC anticonvulsant; and antidepressant. The SESTs can be used for gene  
 CC therapy and in vaccines. The SESTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the SESTs. Proteins encoded by the SESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
 CC in the exemplification of the present invention.  
 CC  
 XX Sequence 922 BP; 243 A; 265 C; 244 G; 166 T; 4 other:  
 SO

Query Match 100.0%; Score 21; DB 21; Length 922;  
 Best Local Similarity 100.0%; Pred. No. 0.3;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 attgtctgagacgctctgctc 21  
 ||||||||||||||||||||  
 DB 118 ATTGCTGTGAGACCGCTGTGTC 98

RESULT 8  
 ID AAS44948 standard; cDNA; 1568 BP.  
 XX  
 AC AAS44948;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE cDNA encoding novel human secretory protein, Seq ID No 29.  
 XX  
 KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;  
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;  
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;  
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;  
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
 KW





XX AAH23086;  
AC 17-SEP-2001 (first entry)  
XX  
DE Osteoarthritis tissue-derived nucleic acid sequence #16.  
XX  
XX Osteoarthritis; infectious disorder; Crohn's disease; sepsis; human;  
KW wound healing; osteopathic; anti-arthritis; anti-inflammatory;  
XX vulnerability; antibacterial; antiallergic; ds.  
XX  
OS Homo sapiens.  
XX WO200153531-A2.  
PN 26-JUL-2001.  
PD 18-JAN-2001; 2001WO-US00016.  
PF 18-JAN-2000; 2000US-0176523.  
XX 18-JAN-2000; 2000US-0176523.  
PR (PHAA ) PHARMACIA CORP.  
PA Phleppard D, Vasanthakamur G, Dotson S, Ma X;  
PI WPI: 2001-451914/48.  
DR  
XX Substantially purified protein, polypeptide or their fragments, used to  
PT identify a biologically active compound or composition and treat  
PT mammalian osteoarthritis -  
XX  
XX Claim 1; Page 102-103; 144pp; English.  
XX  
CC Sequences AAH23071-23152 represent nucleic acid sequences derived from  
CC osteoarthritis tissues. The sequences are useful as probes and for the  
CC diagnosis or prognosis of mammalian osteoarthritis. The polynucleotides  
CC and polypeptides of the invention are useful for generating diagnostic  
CC reagents, as targets for small molecule drug development, generation of  
CC therapeutics, and cloning genes. Specific antibodies are used to generate  
CC enzyme linked immunosorbent assays for detection of osteoarthritis. The  
CC invented molecules can be used to treat osteoarthritis or to analyse the  
CC disease-modifying activity of osteoarthritis drugs. Other disorders  
CC treatable using the nucleic acid sequences include atopic, inflammatory  
CC and infectious disorders e.g. Crohn's disease and sepsis, and wound  
CC healing.  
XX  
SQ Sequence 1678 BP: 424 A; 491 C; 450 G; 313 T; 0 other;  
XX  
Query Match 100.0%; Score 21; DB 22; Length 1678;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 attgtctgagaccgtcgtgctc 21  
DB 156 ATTGCTGTGAGACCGTGTGCTC 136  
AAD05495/C  
ID AAD05495 standard; cDNA; 518 BP.  
AC AAD05495;  
XX  
DT 18-JUL-2001 (first entry)  
XX  
DE Human secreted protein-encoding gene 4 cDNA clone HOPCC3, SEQ ID NO:14.  
XX  
KW Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
XX

KW inflammation; neurological disorder; Alzheimer's disease; food additive;  
KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;  
KW pregnancy-related disorder; endocrine disorder; infection; wound healing;  
KW cell culture; chemotaxis; vulnerability; binding partner identification;  
KW gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Location/Qualifiers  
FT 81..419  
FT CDS  
FT  
FT Key  
FT 81..419  
FT /tag= a  
FT /product= "Human secreted protein precursor"  
FT /transl\_except= (pos:367..369, aa:xaa)  
FT /transl\_except= (pos:414..416, aa:xaa)  
FT /note= "Xaa corresponds to any of the naturally occurring  
FT L-amino acids"  
FT 81..143  
FT sig\_peptide  
FT 144..416  
FT mat\_peptide  
FT /tag= b  
FT /tag= c  
FT /product= "Mature human secreted protein"  
XX  
XX WO200134767-A2.  
XX  
XX 17-MAY-2001.  
XX  
XX 01-NOV-2000; 2000WO-US30036.  
XX  
XX 05-NOV-1999; 99US-0163576.  
XX 27-JUL-2000; 2000US-0221366.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Soppet DR, Komatsoulis G, Shi Y, Olsen HS, Ruben SM;  
PI WPI: 2001-316492/33.  
XX P-PSDB: AAB01675.  
XX  
DR Isolated nucleic acid molecule encoding a human secreted protein is  
XX used in preventing, treating or ameliorating a medical condition -  
XX  
XX Claim 1; Page 413; 540pp; English.  
XX  
CC AAD05492-AAD05564 represent cDNAs corresponding to 22 human secreted  
CC protein genes, and AAB01672-AAB01743 represent the proteins they encode.  
CC AAB01744-AAB01763 represent human secreted protein fragments or variants.  
CC The secreted proteins and their genes are useful for preventing, treating  
CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
CC Pathological conditions can be diagnosed by determining the amount of the  
CC new protein in a sample or by determining the presence of mutations in  
CC the new genes. Specific uses are described for each of the 22 genes, in  
CC based on the tissues in which they are most highly expressed, and include  
CC developing products for the diagnosis or treatment of proliferative  
CC disorders, cancer, tumours, foetal and developmental abnormalities,  
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
CC diseases (e.g., Rheumatoid arthritis), inflammation, allergies,  
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,  
CC pregnancy-related disorders, endocrine disorders, and infections. The  
CC proteins can also be used to aid wound healing and epithelial cell  
CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
CC before transplantation, for supporting cell culture of primary tissues,  
CC to regenerate tissues, to identify their cognate ligands or binding  
CC partners, and in chemotaxis, and can be used as a food additive or  
CC preservative to modify storage properties. Antibodies specific for a  
CC protein of the invention can be used in alleviating symptoms associated  
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
CC The present sequence represents a human secreted protein-encoding cDNA of  
CC the invention.  
XX

SQ Sequence 518 BP; 146 A; 123 C; 148 G; 96 T; 5 other;

Query Match 84.8%; Score 17.8; DB 22; Length 518;  
Best Local Similarity 90.5%; Pred. No. 12;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 atgtctgagaccgtctgctc 21  
1 |||||  
Db 164 AGTGTCTGAGACCATCTGCTC 144

RESULT 12  
AAD05522/c  
ID AAD05522 standard; cDNA: 518 BP.

AC AAD05522;  
XX  
XX 18-JUL-2001 (first entry)

DE Human secreted protein-encoding gene 4 cDNA clone HOF0C33, SEQ ID NO:41.  
XX  
XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
KM foetal abnormality; developmental abnormality; haematopoietic disorder;  
KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
KM Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
KM psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
KM inflammation; neurological disorder; Alzheimer's disease; food additive;  
KM angiotensin disorder; kidney disorder; gastrointestinal disorder; allergy;  
KM pregnancy-related disorder; endocrine disorder; infection; wound healing;  
KM cell culture; chemotaxis; vunerary; binding partner identification;  
KM gene therapy; ss.  
OS  
XX Homo sapiens.

XX  
XX  
FH Key Location/Qualifiers  
FT CDS 81..419  
FT /tag= a  
FT /product= "Human secreted protein precursor"  
FT /transl\_except= (pos:291..293, aa:Xaa)  
FT /transl\_except= (pos:387..389, aa:Xaa)  
FT /transl\_except= (pos:414..416, aa:Xaa)  
FT /note= "Xaa corresponds to any of the naturally occurring  
FT L-amino acids"  
FT 81..131  
FT /tag= b  
FT 132..416  
FT /tag= c  
FT /product= "Mature human secreted protein"

XX  
XX WO200134767-A2.  
XX  
XX 17-MAY-2001.  
XX  
XX 01-NOV-2000; 2000MO-US30036.  
XX  
XX 05-NOV-1999; 99US-0163576.  
XX 27-JUL-2000; 2000US-0221366.  
XX  
XX (HUMA-) HUMAN GENOME SCT INC.  
XX  
XX Sceptet DR, Komatsoulis G, Shi Y, Olsen HS, Ruben SM;  
XX  
XX WPI: 2001-316492/33.  
XX  
XX P-PSDB; AAE01702.  
XX  
XX Isolated nucleic acid molecule encoding a human secreted protein is  
XX used in preventing, treating or ameliorating a medical condition  
XX  
XX Claim 1; Page 434; 540pp; English.  
XX  
XX AAD05492-AAD05564 represent cDNAs corresponding to 22 human secreted  
XX protein genes, and AAE01672-AAE01743 represent the proteins they encode.

CC AAE01744-AAE01763 represent human secreted protein fragments or variants.  
CC The secreted proteins and their genes are useful for preventing, treating  
CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
CC Pathological conditions can be diagnosed by determining the amount of the  
CC new protein in a sample or by determining the presence of mutations in  
CC the new genes. Specific uses are described for each of the 22 genes,  
CC based on the tissues in which they are most highly expressed, and include  
CC developing products for the diagnosis or treatment of proliferative  
CC disorders, cancer, tumours, foetal and developmental abnormalities,  
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
CC angiotensin disorders, kidney disorders, gastrointestinal disorders,  
CC pregnancy-related disorders, endocrine disorders, and infections. The  
CC proteins can also be used to aid wound healing and epithelial cell  
CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
CC before transplantation, for supporting cell culture of primary tissues,  
CC to regenerate tissues, to identify their cognate ligands or binding  
CC partners, and in chemotaxis, and can be used as a food additive or  
CC preservative to modify storage properties. Antibodies specific for a  
CC protein of the invention can be used in alleviating symptoms associated  
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
CC The present sequence represents a human secreted protein-encoding cDNA of  
CC the invention.

SQ Sequence 518 BP; 146 A; 122 C; 148 G; 95 T; 7 other;

Query Match 84.8%; Score 17.8; DB 22; Length 518;  
Best Local Similarity 90.5%; Pred. No. 12;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 atgtctgagaccgtctgctc 21  
1 |||||  
Db 164 AGTGTCTGAGACCATCTGCTC 144

RESULT 13  
AAD08210/c  
ID AAD08210 standard; cDNA: 1670 BP.

AC AAD08210;  
XX  
XX 07-AUG-2001 (first entry)

DE Human secreted protein-encoding gene 1 cDNA clone HOF0C33, SEQ ID NO:30.  
XX  
XX  
XX Human; secreted protein; proliferative disorder; cancer; tumour;  
KM foetal abnormality; developmental abnormality; haematopoietic disorder;  
KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
KM inflammation; allergy; neurological disorder; Alzheimer's disease;  
KM Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
KM skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
KM cardiovascular disorder; angiotensin disorder; kidney disorder;  
KM gastrointestinal disorder; pregnancy-related disorder;  
KM endocrine disorder; infection; wound healing; vunerary;  
KM cell culture; chemotaxis; food additive; gene therapy;  
KM binding partner identification; ss.  
OS  
XX Homo sapiens.

XX  
XX  
FH Key Location/Qualifiers  
FT CDS 76..495  
FT /tag= a  
FT /product= "Human secreted protein precursor"  
FT /transl\_except= (pos:126  
FT 76..126  
FT /tag= b  
FT 127..492  
FT /tag= c  
FT /product= "Mature human secreted protein"

XX XX WO200132837-A1.  
 XX XX 10-MAY-2001.  
 XX XX  
 XX XX 17-OCT-2000; 2000WO-US28664.  
 XX XX  
 XX XX 02-NOV-1999; 99US-0163085.  
 XX XX 17-DEC-1999; 99US-0172411.  
 XX XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Fiacella M, Wei P, Lafleur DM, Olsen HS, Baker K, Ebner R,  
 XX Komatsoulis G, Rosen CA, Ruben SM, Duan RD, Young PE, Florence KA;  
 XX Moore PA, Birse CE, NI J, Soppet DR, Shi Y;  
 XX WPI: 2001-328782/34.  
 XX P-PSDB; AAE03783.  
 XX  
 XX Novel human secreted proteins and nucleic acids for diagnosing,  
 XX preventing and treating neurological, cardiovascular, infectious,  
 XX autoimmune, gastrointestinal, bone disorders, cancer, particularly  
 XX ovarian cancer.  
 XX  
 XX Claim 1: Page 384-385; 421pp; English.  
 XX  
 XX AAD08191-AAD08213 represent cDNAs corresponding to 19 human secreted  
 XX protein genes, and AAE03764-AAE03786 represent the proteins they encode.  
 XX AAE03787-AAE03800 represent human secreted protein fragments or variants.  
 XX The secreted proteins and their genes are useful for preventing,  
 XX treating or ameliorating medical conditions, e.g., by protein or gene  
 XX therapy. Pathological conditions can be diagnosed by determining the  
 XX amount of the new protein in a sample or by determining the presence of  
 XX mutations in the new genes. Specific uses are described for each of the  
 XX 19 genes, based on the tissues in which they are most highly expressed,  
 XX and include developing products for the diagnosis or treatment of  
 XX proliferative disorders, cancer, tumours, foetal and developmental  
 XX abnormalities, hematopoietic disorders, diseases of the immune system,  
 XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,  
 XX allergies, neurological disorders (e.g., Alzheimer's disease,  
 XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,  
 XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,  
 XX cardiovascular disorders, angiogenic disorders, kidney disorders,  
 XX gastrointestinal disorders, pregnancy-related disorders, endocrine  
 XX disorders, and infections. The proteins can also be used to aid wound  
 XX healing and epithelial cell proliferation, to prevent skin aging due to  
 XX sunburn, to maintain organs before transplantation, for supporting cell  
 XX culture of primary tissues, to regenerate tissues, to identify their  
 XX cognate ligands or binding partners, and in chemotaxis, and can be used  
 XX as a food additive or preservative to modify storage properties.  
 XX Antibodies specific for a protein of the invention can be used in  
 XX alleviating symptoms associated with the disorders mentioned above, and  
 XX in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked  
 XX immunosorbent assay (ELISA). The present sequence represents a human  
 XX secreted protein-encoding cDNA of the invention.  
 XX  
 XX Sequence 1670 BP; 425 A; 473 C; 436 G; 336 T; 0 other;  
 XX  
 XX  
 XX Query Match 84.88; Score 17.8; DB 22; Length 1670;  
 XX Best Local Similarity 90.58; Pred. NO.15;  
 XX Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX XX 18-JUL-2001 (first entry)  
 XX XX  
 XX XX Human secreted protein-encoding gene 4 cDNA clone HOPCC33, SEQ ID NO:42.  
 XX DE  
 XX XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
 XX KW foetal abnormality; developmental abnormality; hematopoietic disorder;  
 XX KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 XX KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
 XX KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
 XX KW inflammation; neurological disorder; Alzheimer's disease; food additive;  
 XX KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;  
 XX KW pregnancy-related disorder; endocrine disorder; infection; wound healing;  
 XX KW cell culture; chemotaxis; vulnerability; binding partner identification;  
 XX KW gene therapy; ss.  
 XX  
 XX Homo sapiens.  
 XX OS  
 XX XX  
 XX XX Key Location/Qualifiers  
 XX XX CDS 76..495  
 XX FT /\*tag= a  
 XX FT /product= "Human secreted protein precursor"  
 XX FT sig\_peptide 76..126  
 XX FT /\*tag= b  
 XX FT mat\_peptide 127..492  
 XX FT /\*tag= c  
 XX FT /product= "Mature human secreted protein"  
 XX FT  
 XX XX  
 XX XX WO200134767-A2.  
 XX XX  
 XX XX 17-MAY-2001.  
 XX XX  
 XX XX 01-NOV-2000; 2000WO-US30036.  
 XX XX  
 XX XX 05-NOV-1999; 99US-0163576.  
 XX XX 27-JUL-2000; 2000US-0221366.  
 XX XX  
 XX XX (HUMA-) HUMAN GENOME SCI INC.  
 XX XX  
 XX XX Soppet DR, Komatsoulis G, Shi Y, Olsen HS, Ruben SM;  
 XX WPI: 2001-316492/33.  
 XX P-PSDB; AAE01703.  
 XX  
 XX Isolated nucleic acid molecule encoding a human secreted protein is  
 XX used in preventing, treating or ameliorating a medical condition -  
 XX  
 XX Claim 1: Page 434-435; 540pp; English.  
 XX  
 XX AAD05492-AAD05564 represent cDNAs corresponding to 22 human secreted  
 XX protein genes, and AAE01744-AAE01743 represent the proteins they encode.  
 XX AAE01744-AAE01763 represent human secreted protein fragments or variants.  
 XX The secreted proteins and their genes are useful for preventing, treating  
 XX or ameliorating medical conditions, e.g., by protein or gene therapy.  
 XX Pathological conditions can be diagnosed by determining the amount of the  
 XX new protein in a sample or by determining the presence of mutations in  
 XX the new genes. Specific uses are described for each of the 22 genes,  
 XX based on the tissues in which they are most highly expressed, and include  
 XX developing products for the diagnosis or treatment of proliferative  
 XX disorders, cancer, tumours, foetal and developmental abnormalities,  
 XX hematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 XX diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
 XX neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 XX cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 XX psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 XX angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 XX pregnancy-related disorders, endocrine disorders, and infections. The  
 XX proteins can also be used to aid wound healing and epithelial cell  
 XX proliferation, to prevent skin aging due to sunburn, to maintain organs  
 XX before transplantation, for supporting cell culture of primary tissues,  
 XX to regenerate tissues, to identify their cognate ligands or binding  
 XX partners, and in chemotaxis, and can be used as a food additive or  
 XX preservative to modify storage properties. Antibodies specific for a

CC protein of the invention can be used in alleviating symptoms associated  
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
CC The present sequence represents a human secreted protein-encoding cDNA of  
CC the invention.

XX Sequence 1670 BP; 425 A; 473 C; 436 G; 336 T; 0 other;

Query Match 84.8%; Score 17.8; DB 22; Length 1670;  
Best Local Similarity 90.5%; Pred. No. 15;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 attgtctgagacgctctgctc 21  
1 ||||||||||| |||||||  
Db 159 AGTGTCTGAGACCATCTGCTC 139

RESULT 15  
AAD08191/c  
ID AAD08191 standard; cDNA; 1684 BP.

AC AAD08191;  
XX  
DT 07-AUG-2001 (first entry)

XX Human secreted protein-encoding gene 1 cDNA clone HOFCC33, SEQ ID NO:11.  
DE  
XX  
XX Human secreted protein; proliferative disorder; cancer; tumour;  
KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
KW cardiovascular disorder; angiogenic disorder; kidney disorder;  
KW gastrointestinal disorder; pregnancy-related disorder; gene therapy;  
KW endocrine disorder; infection; wound healing; vulnery;  
KW cell culture; chemotaxis; food additive;  
KW binding partner identification; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH CDS 91..1182  
FT /\*tag= a  
FT /product= "Human secreted protein precursor"  
FT sig\_peptide 91..153  
FT /\*tag= b  
FT mat\_peptide 154..1179  
FT /\*tag= c  
FT /product= "Mature human secreted protein precursor"

PN WO200132837-A1.

XX 10-MAY-2001.

XX 17-OCT-2000; 2000WO-US28664.

XX 02-NOV-1999; 99US-0163085.

PR 17-DEC-1999; 99US-0172411.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Fiscella M, Wei P, Lafleur DW, Olsen HS, Baker K, Ebner R;

PI Komatsoulis G, Rosen CA, Ruben SM, Duan RD, Young PE, Florence KA;

PI Moore PA, Birse CE, Ni J, Soppet DR, Shi Y;

XX WPI: 2001-328782/34.

DR P-PSDB: AAE03764.

PT Novel human secreted proteins and nucleic acids for diagnosing,  
PT preventing and treating neurological, cardiovascular, infectious,  
PT autoimmune, gastrointestinal, bone disorders, cancer, particularly

PT ovarian cancer  
XX  
XX Claim 1; Page 375; 421pp; English.

CC AAD08191-AAD08213 represent cDNAs corresponding to 19 human secreted  
CC protein genes and AAE03764-AAE03787 represent the proteins they encode.  
CC AAE03787-AAE03800 represent human secreted protein fragments or variants.  
CC The genes and their secreted proteins are useful for preventing,  
CC treating or ameliorating medical conditions, e.g., by protein or gene  
CC therapy. Pathological conditions can be diagnosed by determining the  
CC amount of the new protein in a sample or by determining the presence of  
CC mutations in the new genes. Specific uses are described for each of the  
CC 19 genes, based on the tissues in which they are most highly expressed,  
CC and include developing products for the diagnosis or treatment of  
CC proliferative disorders, cancer, tumours, foetal and developmental  
CC abnormalities, haematopoietic disorders, diseases of the immune system,  
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,  
CC allergies, neurological disorders (e.g., Alzheimer's disease,  
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,  
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,  
CC cardiovascular disorders, angiogenic disorders, kidney disorders,  
CC gastrointestinal disorders, pregnancy-related disorders, endocrine  
CC disorders, and infections. The proteins can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues, to identify their  
CC cognate ligands or binding partners, and in chemotaxis, and can be used  
CC as a food additive or preservative to modify storage properties.  
CC Antibodies specific for a protein of the invention can be used in  
CC alleviating symptoms associated with the disorders mentioned above, and  
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked  
CC immunoassay (ELISA). The present sequence represents a human  
CC secreted protein-encoding cDNA of the invention.

XX Sequence 1684 BP; 429 A; 478 C; 435 G; 342 T; 0 other;

Query Match 84.8%; Score 17.8; DB 22; Length 1684;  
Best Local Similarity 90.5%; Pred. No. 15;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 attgtctgagacgctctgctc 21  
1 ||||||||||| |||||||  
Db 174 AGTGTCTGAGACCATCTGCTC 154

Search completed: May 17, 2002, 16:13:55  
Job time: 7047 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:04:13 ; Search time 1959.41 Seconds  
(without alignments)  
224.280 Million cell updates/sec

Title: US-09-913-325-5

Perfect score: 21

Sequence: 1 atgtctgagacgctctgtc 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

C	1	21	100.0	396	6	AX093197	AX093197 Sequence
C	2	21	100.0	1648	9	HUMTRPM2A	M64722 Human TRPM-
C	3	21	100.0	1651	6	A21577	Blood plasm
C	4	21	100.0	1651	9	HUMKLI	M25915 Human compl
C	5	21	100.0	1658	9	BC019588	BC019588 Homo sapi
C	6	21	100.0	1676	9	HSCSP40	X14723 Human SP-40
C	7	21	100.0	1678	6	AX202086	AX202086 Sequence
C	8	21	100.0	1684	9	BC010514	BC010514 Homo sapi
C	9	21	100.0	7610	9	HUMTRPM2A1	M63376 Human TRPM-
C	10	21	100.0	175382	9	AF235104	AF235104 Homo sapi
C	11	21	100.0	187453	2	AC025192	AC025192 Homo sapi
C	12	21	100.0	224431	9	AF311103	AF311103 Homo sapi
C	13	17.8	84.8	136131	9	AC000360	AC000360 Homo sapi
C	14	17.8	84.8	171830	2	AC0092103	AC0092103 Homo sapi
C	15	17.8	84.8	173513	9	HS125H2	Z98949 Human DNA s
C	16	17.8	84.8	191410	10	AC083814	AC083814 Mus muscu
C	17	17.8	84.8	227872	2	AC020786	AC020786 Mus muscu
C	18	17.8	84.8	288888	9	HS010932	HS010932 Homo sapi
C	19	17.4	82.9	177102	9	HS01069C8	HS01069C8 Homo sapi
C	20	16.8	80.0	17345	9	AT062323	AT062323 Homo sapi
C	21	16.8	80.0	43961	1	SCM1	AL133422 Streptomy
C	22	16.8	80.0	120873	9	AL445986	AL445986 Human DNA
C	23	16.8	80.0	138036	2	AP004762	AP004762 Oryza sat
C	24	16.8	80.0	144136	2	AC012243	AC012243 Homo sapi
C	25	16.8	80.0	161547	2	AC093648	AC093648 Homo sapi
C	26	16.8	80.0	163240	2	AC069422	AC069422 Homo sapi
C	27	16.8	80.0	165570	2	AC015765	AC015765 Homo sapi
C	28	16.8	80.0	174952	9	AC002460	AC002460 Human BAC
C	29	16.8	80.0	177664	2	AP004620	AP004620 Oryza sat
C	30	16.8	80.0	179262	2	AC079507	AC079507 Mus muscu
C	31	16.8	80.0	180793	2	AC064812	AC064812 Mus muscu
C	32	16.4	78.1	58431	2	AC100112	AC100112 Mus muscu
C	33	16.4	78.1	58431	2	AC100112	AC100112 Mus muscu
C	34	16.4	78.1	79726	2	AC095616	AC095616 Rattus no
C	35	16.4	78.1	91953	2	AC099305	AC099305 Rattus no
C	36	16.4	78.1	215164	2	AC096975	AC096975 Rattus no
C	37	16.2	77.1	264	6	AX315020	AX315020 Sequence
C	38	16.2	77.1	418	11	G50193	G50193 SHGC-79716
C	39	16.2	77.1	1502	9	AK058077	AK058077 Homo sapi
C	40	16.2	77.1	1678	4	AF118852	AF118852 Oryctolag
C	41	16.2	77.1	1704	9	BC007910	BC007910 Homo sapi
C	42	16.2	77.1	1874	8	AF370145	AF370145 Arabidops
C	43	16.2	77.1	2290	8	ATHANKA	M82883 Arabidops
C	44	16.2	77.1	10381	1	AE007202	AE007202 Simorhizo
C	45	16.2	77.1	11557	1	AE009113	AE009113 Agrobacte

## ALIGNMENTS

RESULT	1	AX093197/c	396 bp	DNA	linear	PAT 30-MAR-2001
LOCUS	AX093197	Sequence 15 from Patent WO0118046.				
DEFINITION	AX093197					
ACCESSION	AX093197.1	GI:13509646				
VERSION						
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE		1 (bases 1 to 396)				
JOURNAL		Xu,J. and Stolk,J.A.				
FEATURES		Ovarian tumor sequences and methods of use therefor				
Source		Patent: WO 0118046-A 15 15-MAR-2001;				
misc-feature		CORIMX CORPORATION (US)				
BASE COUNT		129 a 83 c 121 g 62 t 1 others				

ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 396;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 attgtctgagaccgtctgctc 21  
|||||  
Db 145 ATTGCTGAGACCGTCTGCTC 125

RESULT 2  
LOCUS HUMTRPM2A 1648 bp mRNA linear PRI 23-AUG-1996  
DEFINITION Human TRPM-2 mRNA, complete cds.  
ACCESSION M64722  
VERSION M64722.1 GI:339972  
KEYWORDS TRPM-2 protein.  
SOURCE Human cDNA to mRNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (sites)  
Wong, P., Pineault, J., Lakin, J., Taillefer, D., Leger, J., Wang, C.  
and Tenniswood, M.  
Genomic organization and expression of the rat TRPM-2 (clusterin)  
gene, a gene implicated in apoptosis  
J. Biol. Chem. 268 (7), 5021-5031 (1993)  
MEDLINE 93186813  
REFERENCE 2 (bases 1 to 1648)  
Wong, P., Taillefer, D., Lakin, J., Pineault, J., Chader, G. and  
Tenniswood, M.  
Molecular characterization of human TRPM-2/clusterin, a gene  
associated with sperm maturation, apoptosis and neurodegeneration  
Eur. J. Biochem. 221 (3), 917-925 (1994)  
FEATURES  
JOURNAL MEDLINE 94237156  
BASE COUNT 408 a 488 c 440 g 312 t  
ORIGIN

1. 1648  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1. 1648  
/gene="TRPM-2"  
53. 1402  
/gene="TRPM-2"  
/codon\_start=1  
/protein\_id="AAB06508.1"  
/db\_xref="GI:339973"  
/translation="MMKTLLEFGLLTWESGOVLGDQTVSDNELQEMNSQSKYVK  
EIONAVNGVKQITLKEKTEERKTLISLEAKKKEDALNTPRESPTKLELPGVC  
NETMMALMECKPCLOKOTCMKPYARVCRSGSLVGRLEEFLOINSSPYFMNNGRID  
SILENDROQTHMIDVMDHFSRASSIIDELPDREFTEPDQTYHXIDFSIPHRPHF  
FFPKSRIVASLMPFSPYEPNFAHQPFLEHQAQAMDIHHSAPQHPTEFIRE  
EGDDRTVCREIRHNSTGCLRMKDDCKRETLSDVCSSTNPSQAKLREIDESLOVA  
ERLIRKYNELKSYQWKMNTSLLEQLMEQNMVSRLANLTQGEQYLLVTVASHT  
TSDVPSGVTEVYVKLEFSDPTTVYTPVEVSRKNPKFMETVAEKALQEVKKHREE"  
1627. 1632  
polyA\_signal  
/gene="TRPM-2"  
BASE COUNT 408 a 488 c 440 g 312 t  
ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 1648;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 attgtctgagaccgtctgctc 21  
|||||  
Db 139 ATTGCTGAGACCGTCTGCTC 119

RESULT 3  
MEDLINE A21577/c

LOCUS A21577 1651 bp DNA linear PAT 26-JUL-1994  
DEFINITION blood plasma component having a biological activity of inhibiting  
cytolysis mediated by a cytolytic protein.  
ACCESSION A21577  
VERSION A21577.1 GI:579601  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1651)  
CYTOLYSIS INHIBITOR PROTEINS (CLI) AND DNA SEQUENCES CODING FOR  
SAID PROTEINS  
Patent: WO 9105043-A 3 18-APR-1991;  
JOURNAL MEDLINE 9105043  
FEATURES  
LOCUS A21577 1651 bp DNA linear PRI 27-APR-1993  
DEFINITION Human complement cytolysis inhibitor (CLI) mRNA, complete cds.  
ACCESSION M25915  
VERSION M25915.1 GI:180619  
KEYWORDS blood plasma glycoprotein; complement cytolysis inhibitor;  
secretory glycoprotein; seminal plasma protein; soluble terminal  
complement complex (C5b-9s) component.  
SOURCE Human (adult) liver, cDNA to mRNA, clone CLI-1.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1651)  
Jenne, D.E. and Tschopp, J.  
Molecular structure and functional characterization of a human  
complement cytolysis inhibitor found in blood and seminal plasma:  
identity to sulfated glycoprotein 2, a constituent of rat testis  
fluid  
Proc. Natl. Acad. Sci. U.S.A. 86, 7123-7127 (1989)  
MEDLINE 89386692  
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted  
by D.E.Jenne, 05-JUL-1989.  
FEATURES  
JOURNAL MEDLINE 89386692  
BASE COUNT 405 a 481 c 447 g 318 t  
ORIGIN

1. 1651  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
199. 1545  
/codon\_start=1  
/product="blood plasma component having a biological  
activity of inhibiting cytolysis mediated by a cytolytic  
protein"  
/protein\_id="CAA01560.1"  
/db\_xref="GI:579602"  
/db\_xref="SWISS-PROT:P10909"  
/translation="MMKTLLEFGLLTWESGOVLGDQTVSDNELQEMNSQSKYVK  
EIONAVNGVKQITLKEKTEERKTLISLEAKKKEDALNTPRESPTKLELPGVC  
NETMMALMECKPCLOKOTCMKPYARVCRSGSLVGRLEEFLOINSSPYFMNNGRID  
SILENDROQTHMIDVMDHFSRASSIIDELPDREFTEPDQTYHXIDFSIPHRPHF  
FFPKSRIVASLMPFSPYEPNFAHQPFLEHQAQAMDIHHSAPQHPTEFIRE  
EGDDRTVCREIRHNSTGCLRMKDDCKRETLSDVCSSTNPSQAKLREIDESLOVA  
ERLIRKYNELKSYQWKMNTSLLEQLMEQNMVSRLANLTQGEQYLLVTVASHT  
TSDVPSGVTEVYVKLEFSDPTTVYTPVEVSRKNPKFMETVAEKALQEVKKHREE"  
1627. 1632  
polyA\_signal  
/gene="TRPM-2"  
BASE COUNT 405 a 481 c 447 g 318 t  
ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 1651;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 attgtctgagaccgtctgctc 21  
|||||  
Db 282 ATTGCTGAGACCGTCTGCTC 262

RESULT 4  
LOCUS HUMCLI/c 1651 bp mRNA linear PRI 27-APR-1993  
DEFINITION Human complement cytolysis inhibitor (CLI) mRNA, complete cds.  
ACCESSION M25915  
VERSION M25915.1 GI:180619  
KEYWORDS blood plasma glycoprotein; complement cytolysis inhibitor;  
secretory glycoprotein; seminal plasma protein; soluble terminal  
complement complex (C5b-9s) component.  
SOURCE Human (adult) liver, cDNA to mRNA, clone CLI-1.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1651)  
Jenne, D.E. and Tschopp, J.  
Molecular structure and functional characterization of a human  
complement cytolysis inhibitor found in blood and seminal plasma:  
identity to sulfated glycoprotein 2, a constituent of rat testis  
fluid  
Proc. Natl. Acad. Sci. U.S.A. 86, 7123-7127 (1989)  
MEDLINE 89386692  
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted  
by D.E.Jenne, 05-JUL-1989.  
FEATURES  
JOURNAL MEDLINE 89386692  
BASE COUNT 405 a 481 c 447 g 318 t  
ORIGIN



CDS

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
199..1545
/note="complement cytolysis inhibitor precursor"
/codon_start=1
/protein_id="A0A35692.1"
/db_xref="GI:180620"
/translation="MKTLFLVGLLITWESQVIGDOTVSDNEIQEMSNQSKYVNE
IONAVNGKQIKTLKTEKNEKTLNLEAKKKEDALNETRESETKLELPGVCN
ETMALMECKPCLKQTCMKFYARVCSGSLVGRLEEFLNQSSPEYFMWNGDRID
LLENDROQTHMDVMDHFSRASSIIIDELFODRFTREPQDTYHYLPFSLPHRPHF
FPKSRIVSLMEFSPEYELNFAHMFQPLEMIHBAQAMDIHFSPAFQHPPEFIR
GGDDRTVCREIRHNSTGCLRKMDQDCREILSVDCSTNNPSQAKLRRELDSELOVA
RLTRRYNELLKSYOKMKLNTSLLLEQNEOPNMYSRLANLTQGEDOYLYELVTYASH
SDSDVPSGVTEVYVKLFDSDPITVTYVPEVSRKPKMEYAEKALQETRKRRHEE"
199..261
/feature="complement cytolysis inhibitor signal peptide"
mat_peptide
262..876
/product="complement cytolysis inhibitor a-chain"
mat_peptide
877..1542
/product="complement cytolysis inhibitor b-chain"
BASE COUNT
405 a 481 c 447 g 318 t
ORIGIN
Unreported.

```

Query Match 100.0%; Score 21; DB 9; Length 1651;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgtcagaccgtctgctc 21  
|||||

Db 282 ATTGCTGAGACCGTCTGCTC 262

RESULT 5  
LOCUS BC019588/c 1658 bp mRNA linear PRI 22-JAN-2002  
DEFINITION Homo sapiens, clusterin (complement lysis inhibitor, SP-40, 40, sulfated glycoprotein 2, testosterone-repressed prostate message 2, apolipoprotein J), clone MGC:24903 IMAGE:4915444, mRNA, complete cds.

ACCESSION BC019588  
VERSION BC019588.1 GI:18043614  
KEYWORDS MGC.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1658)  
Strausberg, R.  
Direct Submission  
Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnathure, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 29 Row: m Column: 21

FEATURES  
source

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4502904.  
1..1658  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/db\_xref="locusid:1191"  
/clone="MGC:24903 IMAGE:4915444"  
/tissue\_type="Brain, anaplastic oligodendroglioma with 1p/19q loss"  
/clone\_lib="NCI CGAP\_Brn67"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPOK6"  
49..1398  
/codon\_start=1  
/product="clusterin (complement lysis inhibitor, SP-40, 40, sulfated glycoprotein 2, testosterone-repressed prostate message 2, apolipoprotein J)"  
/protein\_id="AAH19588.1"  
/db\_xref="GI:18043615"  
/translation="MKTLFLVGLLITWESQVIGDOTVSDNEIQEMSNQSKYVNE  
IONAVNGKQIKTLKTEKNEKTLNLEAKKKEDALNETRESETKLELPGVC  
NETMALMECKPCLKQTCMKFYARVCSGSLVGRLEEFLNQSSPEYFMWNGDRID  
LLENDROQTHMDVMDHFSRASSIIIDELFODRFTREPQDTYHYLPFSLPHRPHF  
FPKSRIVSLMEFSPEYELNFAHMFQPLEMIHBAQAMDIHFSPAFQHPPEFIR  
EGDDRTVCREIRHNSTGCLRKMDQDCREILSVDCSTNNPSQAKLRRELDSELOVA  
RLTRRYNELLKSYOKMKLNTSLLLEQNEOPNMYSRLANLTQGEDOYLYELVTYASH  
TSDSDVPSGVTEVYVKLFDSDPITVTYVPEVSRKPKMEYAEKALQETRKRRHEE"

BASE COUNT  
423 a 486 c 439 g 310 t  
ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 1658;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgtcagaccgtctgctc 21  
|||||

Db 135 ATTGCTGAGACCGTCTGCTC 115

RESULT 6  
LOCUS HSCSP40/c 1676 bp mRNA linear PRI 22-MAR-1995  
DEFINITION Human SP-40, 40 mRNA for complement-associated protein SP-40, 40 alpha-1 and beta-1 chain.  
X14723  
X14723.1 GI:30250  
VERSION  
KEYWORDS complement-associated protein; serum protein; SP-40, 40 gene.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1676)  
Kirsbaum, L.  
Direct Submission  
Submitted (17-MAR-1989) Kirsbaum L., The University of Melbourne, The Preclinical Centre, School of Veterinary Science, Parkville Victoria 3052, Australia  
2 (bases 1 to 1676)  
Kirsbaum, L., Sharpe, J.A., Murphy, B., d'Apice, A.J., Classon, B., Hudson, P. and Walker, I.D.  
Molecular cloning and characterization of the novel, human complement-associated protein, SP-40, 40: a link between the complement and reproductive systems  
EMBO J. 8 (3), 711-718 (1989)  
89251601  
Clin. Invest. 81:1858-1864(1988).  
Location/Qualifiers  
1..1676  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

clone="LK (107)"  
 /tissue\_type="liver"  
 /clone\_id="lambda g11"  
 48. .113  
 48. .1397  
 /note="SP-40.40 prepropeptide (AA -22 to 427)"  
 /codon\_start=1  
 /protein\_id="CAA32847.1"  
 /db\_xref="GI:30251"  
 /db\_xref="SWISS-PROT:P10909"  
 /translation="MMKTLLEFVGLLTWESGOVLDTQVSDNELQEMSNQSKYVVK  
 EIQNAVGVQIKTLIEKTEERKTLISNEAKKKEDALNETRESERLKLPEVC  
 NETMALMECKPCIKOTCMKPYARVCRSGSLVQLEFLLNOSPEFYMMNGRID  
 SLENDROOHTMLDVMODHFRASSIIDELEFORFREPQDYHYHLPSPRRRHF  
 FFKSRIVRSILMPSPYEPPLNFHAMQOPLEMIHEAOQANDIHHSAPARHPTETIR  
 EGDDRTVCREIRHNSGTCLRMKDCKCKREITISVDCSTNNPQAKLRRLDESLOVA  
 ERLTRKYNELKSTQMKMNTSSLLLEQNLQFMVSRNLATLQGEQYLRVTYVASH  
 TSDSDVSGVTEVYVKLPDSDPITVTVPEVSRKNPKFMETVAEKALQEVKKHREE"

misc\_feature  
 114. .>114  
 /note="beta-chain"  
 114. .1394  
 /note="SP-40.40 prepeptide (AA 1-427)"  
 729. .1394  
 /product="mature alpha-chain (AA 205-427)"  
 1622. .1627  
 /note="pot. polyA signal"  
 /note="pot. polyA signal"  
 436 a 488 c 437 g 315 t  
 BASE COUNT  
 ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 1676;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 attgtctgagaccgtctgtc 21  
 ||||||||||||||||||||  
 Db 134 ATTGCTGAGACCGCTGTGTC 114

RESULT 7  
 LOCUS AX202086 1678 bp DNA linear PAT 30-AUG-2001  
 DEFINITION Sequence 16 from Patent WO0153531.  
 ACCESSION AX202086  
 VERSION AX202086.1 GI:15391872  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1678)  
 Philpard, D., Vasanthakumari, G., Dotson, S. and Ma, X.J.  
 Osteoarthritis tissue derived nucleic acids, polypeptides, vectors,  
 and cells  
 Patent: WO 0153531-A 16 26-JUL-2001.

JOURNAL Pharmacla Corporation (US)  
 FEATURES  
 source Location/Qualifiers  
 1. .1678  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 424 a 491 c 450 g 313 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 1676;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 attgtctgagaccgtctgtc 21  
 ||||||||||||||||||||  
 Db 156 ATTGCTGAGACCGCTGTGTC 136

RESULT 8  
 BC010514/c  
 LOCUS  
 DEFINITION Homo sapiens, clusterin (complement lysis inhibitor, SP-40/40,  
 sulfated glycoprotein 2, testosterone-repressed prostate message 2,  
 apolipoprotein J), clone MGC:18080 IMAGE:4150452, mRNA, complete  
 cds

ACCESSION BC010514  
 VERSION BC010514.1 GI:14714740  
 KEYWORDS MGC.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1684)  
 Strausberg, R.  
 Direct Submission  
 Submitted (10-JUL-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: David N. Louis, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [villalob@bcm.tmc.edu](mailto:villalob@bcm.tmc.edu)  
 Villalob, D.K., Luna, R.A., Hale, S.M., Huiyk, S., Lu, X., Garcia,  
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
 Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAC Plates: 19 Row: 0 Column: 12.  
 Location/Qualifiers

1. .1684  
 /organism="Homo sapiens"  
 /db\_xref="locusID:1191"  
 /db\_xref="taxon:9606"  
 /clone="MGC:18080 IMAGE:4150452"  
 /tissue\_type="Brain, glioblastoma with EGFR amplification"  
 /clone\_id="NCI CGAP\_Brn64"  
 /lab\_host="DHI0B"  
 /note="Vector: pCMV-SPORT6"  
 69. .1418  
 /codon\_start=1  
 /product="clusterin (complement lysis inhibitor, SP-40/40,  
 sulfated glycoprotein 2, testosterone-repressed prostate  
 message 2, apolipoprotein J)"  
 /protein\_id="AAH10514.1"  
 /db\_xref="GI:14714741"

/translation="MMKTLLEFVGLLTWESGOVLDTQVSDNELQEMSNQSKYVVK  
 EIQNAVGVQIKTLIEKTEERKTLISNEAKKKEDALNETRESERLKLPEVC  
 NETMALMECKPCIKOTCMKPYARVCRSGSLVQLEFLLNOSPEFYMMNGRID  
 SLENDROOHTMLDVMODHFRASSIIDELEFORFREPQDYHYHLPSPRRRHF  
 FFKSRIVRSILMPSPYEPPLNFHAMQOPLEMIHEAOQANDIHHSAPARHPTETIR  
 EGDDRTVCREIRHNSGTCLRMKDCKCKREITISVDCSTNNPQAKLRRLDESLOVA  
 ERLTRKYNELKSTQMKMNTSSLLLEQNLQFMVSRNLATLQGEQYLRVTYVASH  
 TSDSDVSGVTEVYVKLPDSDPITVTVPEVSRKNPKFMETVAEKALQEVKKHREE"  
 BASE COUNT 433 a 493 c 448 g 310 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 1684;  
 Best Local Similarity 100.0%; Pred. No. 1.2;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 attgtctgagaccgtctgtc 21

```
|||||
Db      155 ATTGCTGAGACCGCTGTGTC 135

RESULT 9
LOCUS   HUMTRPM2A1/c
DEFINITION Human TRPM-2 protein gene, exons 1,2 and 3.
ACCESSION M63376
VERSION  M63376.1 GI:292838
KEYWORDS TRPM-2 protein.
SEGMENT 1 of 4
SOURCE   Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Wong,P., Pineault,J., Lakin,J., Taillefer,D., Leger,J., Wang,C.
and Tenniswood,M.
TITLE    Genetic organization and expression of the rat TRPM-2 (clusterin)
JOURNAL  J. Biol. Chem. 268 (7), 5021-5031 (1993)
MEDLINE  9316813
REFERENCE
AUTHORS Wong,P., Taillefer,D., Lakin,J., Pineault,J., Chader,G. and
Tenniswood,M.
TITLE    Molecular characterization of human TRPM-2/clusterin, a gene
JOURNAL  Eur. J. Biochem. 221 (3), 917-925 (1994)
MEDLINE  94237156
FEATURES
source   1..7610
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         1422..1469
         /gene="TRPM-2"
         /number=1
         5509..5634
         /gene="TRPM-2"
         /number=2
         7021..7169
         /gene="TRPM-2"
         /number=3
BASE COUNT 1660 a 1994 c 2055 g 1899 t 2 others
ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 7610;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtctgagaccgtctgctc 21
|||||
Db 5624 ATTGCTGAGACCGCTGTGTC 5604

RESULT 10
LOCUS   AF235104
DEFINITION Homo sapiens chromosome 8 clone GSI-251107 map 8p12, complete
sequence.
ACCESSION AF235104
VERSION  AF235104.3 GI:16543385
KEYWORDS HTG: HTGS_FULFITOR; HTGS_ACTIVEFIN.
SOURCE   human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Men,G., Reichwald,K., Baumgart,C., Taudin,S., Baumgart,C. and
Platzer,M.
TITLE    Chromosome 8 genomic sequence
JOURNAL  Unpublished
```

```
REFERENCE
AUTHORS Reichwald,K., Menzel,U., Dette,M., Baumgart,C., Jahn,N., Men,G.,
Schillhabel,M. and Kosenenthal,A.
TITLE    Direct Submission
JOURNAL  Submitted (16-FEB-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE
AUTHORS 3 (bases 1 to 175382)
TITLE    Genome Sequencing Center Jena.
JOURNAL  Direct Submission
REFERENCE
AUTHORS Submitted (24-MAY-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
TITLE    4 (bases 1 to 175382)
AUTHORS Men,G. and Platzer,M.
JOURNAL  Direct Submission
REFERENCE
AUTHORS Submitted (06-FEB-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
COMMENT  On Feb 6, 2002 this sequence version replaced gi:8152056.
-----
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
-----
Project Information
Center project name: H342
Center clone name: GSI-251107
-----
Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 172785 bases at least Q40
Consensus quality: 174289 bases at least Q30
Consensus quality: 174952 bases at least Q20
Quality coverage: 6.88 x in Q20 bases; sum-of-contigs
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the phrap assembly program.
All manually edited bases have been reduced to quality 10.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.
-----
Location/Qualifiers
source   1..175382
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /chromosome="8"
         /map="8p12"
         /clone="GSI-251107"
BASE COUNT 46563 a 38626 c 38604 g 51589 t
ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 175382;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtctgagaccgtctgctc 21
|||||
Db 173484 ATTGCTGAGACCGCTGTGTC 173464

RESULT 11
LOCUS   AC025192/c
DEFINITION Homo sapiens chromosome 8 clone RP11-509E2 map 8, WORKING DRAFT
SEQUENCE, 7 unordered pieces.
ACCESSION AC025192
VERSION  AC025192.3 GI:12313839
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE   human.
```

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 187453)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome 8, clone RP11-509E2  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 187453)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,  
Boguslavsky, L., Bouckhalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J., S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardina, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J., C., Iliev, I., Johnson, R., Jones, C., Kan, L., Karatas, A.,  
Klein, J., Larocque, K., Lamazares, R., Landers, T., Lechoczy, J.,  
Levine, R., Liu, G., Liu, G., Locke, K., MacDonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,  
Meidrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C., H., O'Connor, T., O'Donnell, P.,  
O'Neil, D., Oliver, T., M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Teafaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE Direct Submission  
JOURNAL Submitted (07-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT All repeats were identified using RepeatMasker:  
Smit, A.P.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center code: W1BR  
Center: Whitehead Institute/ MIT Center for Genome Research  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information  
Center project name: L5592  
Center clone name: 509\_E2

Summary Statistics  
Sequencing vector: M13; M77815; 48% of reads  
Sequencing vector: Plasmid; n/a; 52% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 185120 bases at least Q40  
Consensus quality: 186533 bases at least Q20  
Insert size: 175000; agarose-fp  
Insert size: 186853; sum-of-ctrls  
Quality coverage: 8.3 in Q20 ba.  
Quality coverage: 8.3 in Q20 ba.

NOTE: This is a 'working draft' sequence. It currently  
consists of 7 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 17773: contig of 17773 bp in length  
17774 17873: gap of 100 bp  
17874 20278: contig of 2405 bp in length  
20279 20378: gap of 100 bp  
20379 35742: contig of 15364 bp in length  
35743 35842: gap of 100 bp  
35843 65413: contig of 29571 bp in length  
65414 65513: gap of 100 bp  
65514 112793: contig of 47280 bp in length

FEATURES  
source  
1..187453  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-509E2"  
/clone\_lib="RP11-509E2 Human Male BAC"  
1..17773  
/note="assembly\_fragment"  
clone\_end:sp6  
vector\_side:left  
17874..20278  
/note="assembly\_fragment"  
20379..35742  
/note="assembly\_fragment"  
35843..65413  
/note="assembly\_fragment"  
65514..112793  
/note="assembly\_fragment"  
112894..150031  
/note="assembly\_fragment"  
150132..187453  
/note="assembly\_fragment"  
clone\_end:17  
vector\_side:right

BASE COUNT 51878 a 43372 c 44058 g 47545 t 600 others  
ORIGIN

Query Match 100.0%; Score 21; DB 2; Length 187453;  
Best Local Similarity 100.0%; Pred. NO. 0.82;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 atgtctgagacgtctgtc 21  
|||||  
Db 124941 ATGTCTGAGACGTCGTGCTC 124921

RESULT 12  
AF311103/C  
LOCUS Homo sapiens chromosome 8 clone SCB-212e3 map 8p12, complete  
DEFINITION  
ACCESSION AF311103  
VERSION AF311103.3 GI:18542957  
KEYWORDS HTG; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 234431)  
Wen, G., Blechschmidt, K., Baumgart, C., Taudin, S., Baumgart, C. and  
Platzer, M.

REFERENCE 1  
AUTHORS Chromosome 8 genomic sequence  
JOURNAL Unpublished  
TITLE 2 (bases 1 to 234431)  
AUTHORS Blechschmidt, K., Jandrig, B., Baumgart, C., Detle, M., D., Jahn, N.,  
Menzel, U., Schilhabel, M.B., Wen, G., Taudien, S. and Rosenthal, A.  
JOURNAL Direct Submission  
TITLE Submitted (04-OCT-2000) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
3 (bases 1 to 234431)  
Wen, G. and Platzer, M.

REFERENCE 2  
AUTHORS Direct Submission  
TITLE Submitted (06-FEB-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany  
COMMENT On Feb 6, 2002 this sequence version replaced gi:14389409.  
----- Genome Center

Center: Institute of Molecular Biotechnology  
 Center code: IMB  
 Web site: <http://genome.imb-jena.de/>  
 Contact: [gscj-submit@genome.imb-jena.de](mailto:gscj-submit@genome.imb-jena.de)  
 ----- Project Information -----  
 Center project name: H319  
 Center clone name: SCB-212e3  
 ----- Summary Statistics -----  
 Sequencing vector: M13: 100% of reads  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 233404 bases at least Q40  
 Consensus quality: 234413 bases at least Q30  
 Consensus quality: 234428 bases at least Q20  
 Quality coverage: 5.81 x in Q20 bases; sum-of-contrigs

----- Sequence Quality Assessment -----  
 This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

# FEATURES

Source Location/Qualifiers  
 1..234431  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8p12"  
 /clone="SCB-212e3"  
 BASE COUNT 62993 a 56257 c 55890 g 59291 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 234431;  
 Best Local Similarity 100.0%; Pred. No. 0.8;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 attgtctgagacgctctgctc 21  
 ||||||||||||||||||||  
 Db 14062 ATTGCTGAGACGCTCTGCTC 14042

RESULT 13  
 AC000360 136131 bp DNA linear PRI 03-JAN-2001  
 LOCUS Homo sapiens Chromosome 2 BAC Clone 376a1, complete sequence.  
 DEFINITION AC000360  
 ACCESSION AC000360 GI:12018422  
 VERSION  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 136131)  
 AUTHORS McDermid,H.E.  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 136131)  
 AUTHORS Kim,D.-J. and Simon,M.I.  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 136131)  
 AUTHORS Crabtree,J.S. and Roe,B.A.  
 TITLE Homo sapiens Chromosome 2 BAC Clone 376a1  
 JOURNAL Unpublished  
 REFERENCE 4 (bases 1 to 136131)  
 AUTHORS Lantui,L. and Roe,B.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-MAR-1997) Department Of Chemistry And Biochemistry,  
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA

REFERENCE 5 (bases 1 to 136131)  
 AUTHORS Crabtree,J.S., McDermid,H.E., Kim,U.-J., Simon,M.I. and Roe,B.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-JAN-2001) Department Of Chemistry And Biochemistry,  
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
 OK 73019, USA

COMMENT On Jan 3, 2001 this sequence version replaced gi:12000483.  
 FEATURES location/Qualifiers  
 source 1..136131  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /clone="376a1"

BASE COUNT 37449 a 29927 c 29780 g 38975 t  
 ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 136131;  
 Best Local Similarity 90.5%; Pred. No. 50;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 attgtctgagacgctctgctc 21  
 ||||||||||||||||||||  
 Db 62757 ATTGCTGAGACGCTCTGCTC 62777

RESULT 14  
 AC092103 171830 bp DNA linear HTG 24-JAN-2002  
 LOCUS Homo sapiens chromosome 2 clone RP11-143D11, WORKING DRAFT  
 DEFINITION AC092103  
 ACCESSION AC092103  
 VERSION AC092103 GI:18308811  
 KEYWORDS HTGS-PHASE1; HTGS-DRAFT; HTGS-FULLTOP.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 171830)  
 AUTHORS Waterston,R.H.  
 TITLE The sequence of Homo sapiens clone  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 171830)  
 AUTHORS Waterston,R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-JUN-2001) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 COMMENT On Jan 24, 2002 this sequence version replaced gi:16799048.

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc/index.shtml>  
 Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
 ----- Project Information -----  
 Center project name: H.NH0143D11  
 ----- Summary Statistics -----  
 Sequencing vector: M13: 0%  
 Sequencing vector: plasmid: 100%  
 Chemistry: Dye-Primer ET: 0% of reads  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap: version 0.990319  
 Consensus quality: 168799 bases at least Q40  
 Consensus quality: 168467 bases at least Q30  
 Consensus quality: 169732 bases at least Q20  
 Insert size: 17000; agarose-1p  
 Insert size: 170339; sum-of-contrigs  
 Quality coverage: 8.19 in Q20 bases; agarose-1p  
 Quality coverage: 7.96 in Q20 bases; sum-of-contrigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently

\* consists of 12 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1      1149: contig of 1149 bp in length
*      1249: gap of unknown length
*      1250      2446: contig of 1197 bp in length
*      2447      2546: gap of unknown length
*      2547      2937: contig of 391 bp in length
*      2938      3037: gap of unknown length
*      3038      23810: contig of 20773 bp in length
*      23811      23910: gap of unknown length
*      23911      27801: contig of 3891 bp in length
*      27802      32273: gap of unknown length
*      32274      32373: contig of 4372 bp in length
*      32374      37772: gap of unknown length
*      37773      37872: gap of 5399 bp in length
*      37873      51582: contig of 13710 bp in length
*      51583      51682: gap of unknown length
*      51683      67875: contig of 16193 bp in length
*      67876      67975: gap of unknown length
*      67976      86364: contig of 18389 bp in length
*      86365      86464: gap of unknown length
*      86465      104664: contig of 18200 bp in length
*      104665      104764: gap of unknown length
*      104765      171830: contig of 67066 bp in length.
  
```

FEATURES

```

source
1. 171830
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="2"
   /clone="RP11-143D11"
   1. 1149
      misc_feature
      /note="assembly_name:Contig15"
      1250. 2446
      misc_feature
      /note="assembly_name:Contig28"
      2547. 2937
      misc_feature
      /note="assembly_name:Contig41"
      3038. 23810
      /note="assembly_name:Contig62"
      clone_end:17
      vector_side:left"
      23911. 27801
      misc_feature
      /note="assembly_name:Contig55"
      27902. 32273
      misc_feature
      /note="assembly_name:Contig56"
      32374. 37772
      misc_feature
      /note="assembly_name:Contig57"
      37873. 51582
      misc_feature
      /note="assembly_name:Contig58"
      51683. 67875
      misc_feature
      /note="assembly_name:Contig59"
      67976. 86364
      misc_feature
      /note="assembly_name:Contig60"
      86465. 104664
      /note="assembly_name:Contig61"
      104765. 171830
      misc_feature
      /note="assembly_name:Contig63"
BASE COUNT 47788 a 37462 c 36363 g 49114 t 1103 others
ORIGIN

```

Query Match 84.88: Score 17.8; DB 2; Length 171830;  
 Best Local Similarity 90.5%; Pred. No. 49;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 attgtcagagcgtctgac 21  
 |||||||  
 Db 122060 ATTGTGAGACATTTCTGTC 122080

RESULT 15  
 HS125H2 173513 bp DNA linear PRI 12-DEC-1999  
 LOCUS  
 DEFINITION Human DNA sequence from clone CTA-125H2 on chromosome 22q11.12  
 Contains part of the gene for a novel protein similar to KIAA0216  
 and myosin heavy chain, ESTs, GSSs and a Cpg island, complete  
 sequence.

ACCESSION Z98949.1  
 VERSION Z98949.1 GI:3168990  
 KEYWORDS HTG; Cpg island; KIAA0216; myosin.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 173513)

REFERENCE  
 AUTHORS Lloyd, P.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 requests: clonerequests@sanger.ac.uk

COMMENT  
 On May 30, 1998 this sequence version replaced gi:2558540.  
 This sequence has been finished according to sequence map criteria  
 as follows. An attempt is made to resolve all sequencing problems,  
 such as compressions and repeats, but not necessarily within known  
 annotated human repeat sequence elements (e.g. Alu). Where the  
 sequence is ambiguous, there is an annotation using the 'unsure'  
 feature key.

This sequence was generated from part of bacterial clone contigs of  
 human chromosome 22, constructed by the Sanger Centre Chromosome 22  
 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr22  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information  
 on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep  
 CTA-125H2 is from the human BAC library described in U-J. Kim et al. (1996)  
 Genomics 34, 213-218.  
 VECTOR: pBAC108L  
 This sequence is the entire insert of clone CTA-125H2 The start of  
 this sequence overlaps with sequence AL002329 The end of this  
 sequence overlaps with sequence AL000000.  
 Location/Qualifiers  
 1. 173513  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="22"  
 /map="q11-12"  
 /clone="CTA-125H2"  
 /clone\_id="CT978SK-A1"  
 231..371  
 repeat\_region  
 gene  
 /note="WIR repeat: matches 40..192 of consensus"  
 /gene="DK125H2.1"  
 157629..157759  
 /gene="DK125H2.1"  
 16737..16846,33587..33652,39553..39681,41456..41599,  
 43226..43315,56023..56142,56571..56679,60423..60547,  
 63574..63728,67880..68003,68898..69098,73589..73699,  
 76213..76317,86524..86676,111403..111516,112978..113094,  
 115633..115716,117513..117669,120445..120630,  
 157629..157759)  
 157629..157759)  
 16737..16846,33587..33652,39553..39681,41456..41599,  
 43226..43315,56023..56142,56571..56679,60423..60547,  
 63574..63728,67880..68003,68898..69098,73589..73699,  
 76213..76317,86524..86676,111403..111516,112978..113094,

115633. .115716,117513. .117669,120445. .120630,  
157629. >157759)  
/gene="bk125H2.1"  
/note="match: CDNAS: Em:D86970  
match: ESTs: Em:AA504934 Em:R41006  
match: proteins: Tr:060772 Tr:092614  
match: genomic DNA: Em:H55072 Em:H55541 Em:H55303"  
/codon\_start=1  
/evidence=not experimental  
/product="bk125H2.1 (novel protein similar to KIA0216 and  
myosin heavy chain)"  
/protein\_id="CAB11606.2"  
/db\_xref="SPTREMBL:060772"  
/translation="VRLPAGGAGADARGLFWLDEEVHVGSSDVLRLCAAEFK  
GAGTGGSSALRTCEQCELPCEHOLGMDPVRYDTGLHFAKPNLSALAPOLHOSK  
RRELRLFOARKLPPVCAVAGLPGTSOALORSRVRRTPSLGARRAPGCSOI  
KIOMALTSMKRSLRTHICTYPNPVYESGSGESPPOPRDRPGAGPLADIP  
ALRVOLAGPHILRLRHTGADHMGILFRROQVLDAPLKLKSTSEGLDERKA  
VEELLETLDLEKKAVALGSHQVFLKAGVISRLERKQKLSQSVLFLQAACKGFSRQ  
EERKLRIRLAOCIQKNVAFLVKDPMWQLGSLQPLSATIGTEOLRAKEEELT  
TLRRLEKSEKLRNELRQNTDLESKIADLSLADERFEGVACVLESEARLOA  
PREVOELSKSEHQKRLGDYVKOLEPAQOKIQLNDLERPFGADEWQRFCAOME  
KEKHKHLOGOGMSNRILSYLQLOVAKMREYLEOSTVRAJVSROEAVICLEKRT  
ERQKVOIKRFEVLVIRLDSILKMGESLSQATRETSQORSSQYTORLEEKLEDEE  
LVQREAEASRRCMLEKYEVEELAAVRQTLQDLETSTRADLAALAEVASSDSOTE  
"  
855. .924  
/note="L2 repeat: matches 2686. .2749 of consensus"  
1660. .1713  
/note="18 copies 3 mer tga 72 conserved"  
1946. .2233  
/note="AluSq repeat: matches 1. .287 of consensus"  
2234. .2278  
/note="15 copies 3 mer taa 100 conserved"  
2362. .2423  
/note="L2 repeat: matches 2679. .2741 of consensus"  
2489. .2836  
/note="L2 repeat: matches 1257. .1637 of consensus"  
3136. .3334  
/note="L1PA11 repeat: matches 5960. .6165 of consensus"  
3384. .3475  
/note="23 copies 4 mer atat 72 conserved"  
3392. .3473  
/note="41 copies 2 mer at 76 conserved"  
3483. .3743  
/note="AluSq repeat: matches 24. .283 of consensus"  
3750. .4067  
/note="MER7A repeat: matches 1. .329 of consensus"  
4073. .4252  
/note="45 copies 4 mer tgtg 73 conserved"  
4077. .4250  
/note="87 copies 2 mer tg 73 conserved"  
4253. .4552  
/note="75 copies 4 mer gtgt 75 conserved"  
4555. .4640  
/note="ORSL repeat: matches 189. .274 of consensus"  
5228. .5472  
/note="L2 repeat: matches 2434. .2693 of consensus"  
5497. .5647  
/note="L2 repeat: matches 2286. .2442 of consensus"  
7015. .7116  
/note="L2 repeat: matches 2561. .2669 of consensus"  
7200. .7239  
/note="L2 repeat: matches 2580. .2619 of consensus"  
7658. .7961

/note="AluSq repeat: matches 1. .305 of consensus"  
8081. .8163  
/note="L2 repeat: matches 2619. .2710 of consensus"  
8230. .8419  
/note="MIR repeat: matches 57. .245 of consensus"  
complement(9497. .9973)  
/note="match: GSS: Em:AQ00189"  
9769. .9950  
/note="MIR repeat: matches 54. .250 of consensus"  
10305. .10353  
/note="MIR repeat: matches 97. .149 of consensus"  
10504. .10811  
/note="AluSc repeat: matches 1. .308 of consensus"  
11638. .11776  
/note="MIR repeat: matches 46. .188 of consensus"  
12067. .12375  
/note="AluSc repeat: matches 1. .309 of consensus"  
12426. .12515  
/note="MIR repeat: matches 47. .136 of consensus"  
13064. .13180  
/note="MTR repeat: matches 20. .134 of consensus"  
complement(13160. .13420)  
/note="match: GSS: Em:AQ009382"  
13204. .13243  
/note="10 copies 4 mer acac 95 conserved"  
13246. .13314  
/note="L1PA7 repeat: matches 6057. .6125 of consensus"  
13444. .13868  
/gene="bk125H2.1"  
/note="match: GSS: Em:AQ060882"  
13447. .13881  
/gene="bk125H2.1"  
/note="match: GSS: Em:AQ409072"  
13449. .13908  
/gene="bk125H2.1"  
/note="match: GSS: Em:AQ819687"  
13535. .13778  
/note="L2 repeat: matches 2273. .2554 of consensus"  
14793. .14826  
/note="L2 repeat: matches 2695. .2728 of consensus"  
14901. .15097  
/note="MIR repeat: matches 4. .221 of consensus"  
15155. .15308  
/note="MIR repeat: matches 31. .190 of consensus"  
15636. .15831  
/note="MIR repeat: matches 44. .262 of consensus"  
15848. .16027  
/note="AluJb repeat: matches 112. .297 of consensus"  
16056. .16185  
/note="L2 repeat: matches 2355. .2487 of consensus"  
16918. .17021  
/note="AluJ/FLAM repeat: matches 2. .101 of consensus"  
17444. .17555  
/note="L2 repeat: matches 2363. .2475 of consensus"  
17623. .17669  
/note="L2 repeat: matches 2665. .2710 of consensus"  
17860. .17971  
/note="MIR repeat: matches 77. .192 of consensus"  
  
Query Match 84.8%; Score 17.8; DB 9; Length 173513;  
Best Local Similarity 90.5%; Pred. No. 49;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 attgtctagacgctgctgctc 21  
|||||  
Db 167590 ATTGCTGAGACCTTCTGCTC 167610  
  
Search completed: May 17, 2002, 16:04:32  
Job time: 16132 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 15:25:38 ; Search time 3628.41 Seconds  
(without alignments)  
78.116 Million cell updates/sec

Title: US-09-913-325-12

Perfect score: 21  
Sequence: 1 gctggcgagctgggggacct 21

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estrpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_estl2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	121	10	H84500 yv02d12.s1
2	21	100.0	121	10	R93611 yq38d11.s1
3	21	100.0	135	10	T53781 yb33e02.s1
4	21	100.0	143	10	BE765434 IL3-NT010
5	21	100.0	145	9	AM904670 RCI-NN106
6	21	100.0	146	9	AA738045 nx15f08.s
7	21	100.0	151	9	AA434384 zw31h11.s
8	21	100.0	151	10	B1055204 PM2-GN049
9	21	100.0	153	10	B1040715 PM2-NT023
10	21	100.0	157	10	B1040715 PM2-NT023
11	21	100.0	159	10	B1040516 PM2-NT023
12	21	100.0	169	9	A1523434 ar72h08.x
13	21	100.0	169	9	A1691105 tx90g01.x
14	21	100.0	171	10	BF897129 IL0-NT021
15	21	100.0	171	10	BF919288 IL0-NT017
16	21	100.0	172	9	AA235650 z536f04.s
17	21	100.0	172	10	BF932077 PM2-NT023

Result	Score	Query Match	Length	DB ID	Description
18	21	100.0	174	9	AA635704 n33e02.s
19	21	100.0	174	10	BE702836
20	21	100.0	174	10	BF374244 MRL-SN006
21	21	100.0	176	10	B1040542 PM2-NT023
22	21	100.0	176	10	BF364496 QV2-NN107
23	21	100.0	177	10	B1040538 PM2-NT023
24	21	100.0	179	9	AM868593 MRL-SN006
25	21	100.0	179	10	B1040710 PM2-NT023
26	21	100.0	181	10	B1040535 PM2-NT023
27	21	100.0	182	9	AM868587 MRL-SN006
28	21	100.0	184	9	AA433884 zw31h11.r
29	21	100.0	185	10	R83374 ypo6f05.s1
30	21	100.0	187	9	AM897945 RCI-NN006
31	21	100.0	187	10	BE765800 IL3-NT010
32	21	100.0	187	10	BE766185 IL3-NT010
33	21	100.0	190	10	BF749857 CM3-BN038
34	21	100.0	193	10	B1040729 PM2-NT023
35	21	100.0	195	9	AM904698 RCI-NN106
36	21	100.0	196	10	BF749861 CM3-BN038
37	21	100.0	196	10	B1040731 PM2-NT023
38	21	100.0	196	10	B1055202 PM2-NT049
39	21	100.0	198	9	AM243998 x448e10.x
40	21	100.0	198	9	AM868592 MRL-SN006
41	21	100.0	199	9	AM904691 RCI-NN106
42	21	100.0	199	10	BF823349 PM2-RT005
43	21	100.0	200	10	B1055200 PM2-GN049
44	21	100.0	201	9	AM897949 RCI-NN006
45	21	100.0	201	10	B1040531 PM2-NT023

## ALIGNMENTS

RESULT 1  
H84500 LOCUS  
DEFINITION yv02d12.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone IMAGE:241559 3' similar to gb:X14723 CLUSTERIN PRECURSOR (HUMAN); mRNA sequence.  
H84500  
H84500.1 GI:1063171  
EST.

## ACCESSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

/clone.lib="Soares fetal liver spleen INFLS"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: Liver and Spleen; Vector: pRT3D (Pharmacia)  
with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Pac I oligo(dT) primer  
(5' AACCTGACAGATTAAATTAAGATCTTTTCTTTTCTT 3'),  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pRT3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Felina Bonaldo."

BASE COUNT 27 a 20 c 45 g 20 t 9 others

Query Match 100.0%; Score 21; DB 10; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2, 1e+02;  
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctggcgaggttggtggcct 21  
|||||  
Db 95 GCTGGCGAGCTGGGCGCT 115

RESULT 2  
R93611 121 bp mRNA linear EST 29-ANG-1995  
LOCUS y936d11.s1 Soares fetal liver spleen INFLS cDNA clone  
DEFINITION IMAGE:198069 3' similar to gb:X14723 CLUSTERIN PRECURSOR (HUMAN);  
mRNA sequence.  
ACCESSION R93611.1 GI:967777  
VERSION EST.  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
1 (bases 1 to 121)  
Hillier, L., Clackson, N., Dubucque, T., Elliston, K., Hawkins, M., Holman  
, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Merri, M., Parsons, J.,  
Ritkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston  
, R., Williamson, A., Wohlmann, P., and Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)  
Contact: Wilson R  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 1368  
High quality sequence starts: 1  
High quality sequence stops: 1  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Trace considered overall poor quality  
Insert Length: 1368 Std Error: 0.00  
Seq primer: Promega -21ml3  
High quality sequence stop: 1.  
Location/Qualifiers  
1. 121  
/organism="Homo sapiens"  
/db\_xref="GDB:3767119"  
/db\_xref="taxon:9606"  
/clone="IMAGE:198069"  
/clone.lib="Soares fetal liver spleen INFLS"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: Liver and Spleen; Vector: pRT3D (Pharmacia)

with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
(5' AACCTGACAGATTAAATTAAGATCTTTTCTTTTCTT 3'),  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pRT3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Felina Bonaldo."

BASE COUNT 23 a 22 c 50 g 23 t 3 others

Query Match 100.0%; Score 21; DB 10; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2, 1e+02;  
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctggcgaggttggtggcct 21  
|||||  
Db 85 GCTGGCGAGCTGGGCGCT 105

RESULT 3  
T53781 135 bp mRNA linear EST 06-FEB-1995  
LOCUS yb83e02.s1 Stratagene liver (#937224) Homo sapiens cDNA clone  
DEFINITION IMAGE:77786 3' similar to gb:X14723 CLUSTERIN PRECURSOR  
(HUMAN), mRNA sequence.  
ACCESSION T53781.1 GI:655642  
VERSION EST.  
KEYWORDS human.  
SOURCE  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
1 (bases 1 to 135)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chappelli, B.,  
Christoe, S., Dietrich, N., Dubucque, T., Favell, A., Gish, M., Hawkins  
, B., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore  
, B., Morris, M., Parsons, J., Prange, C., Ritkin, L., Rohlfing, T.,  
Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevasakis, E.,  
Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Merri, M.  
Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)  
97044478  
Contact: Wilson R  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 1905  
High quality sequence starts: 1  
High quality sequence stops: 1  
Source: IMAGE Consortium, LNL; This clone is available royalty-free  
through LNL; contact the IMAGE Consortium (info@image.lnl.gov)  
for further information. Trace considered overall poor quality  
Insert Length: 1905 Std Error: 0.00  
Seq primer: -21ml3  
High quality sequence stop: 1.  
Location/Qualifiers  
1. 135  
/organism="Homo sapiens"  
/db\_xref="GDB:497531"  
/db\_xref="taxon:9606"  
/clone="IMAGE:77786"  
/clone.lib="Stratagene liver (#937224)"  
/sex="male"  
/dev\_stage="49 years old"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/note="Organ: liver; Vector: pBluescript SK; Site\_1: EcoRI  
; Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo  
dT. Hepatectomy from normal male caucasian. Average insert  
size: 1.1 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'  
GAATTCGGCAGAG 3' -3' adaptor sequence: 5'



VERSION AA738045.1 GI:2768802  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 146)  
 AUTHORS NCI-CCAP http://www.nci.nlm.nih.gov/nciccap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bcrfemail.nih.gov  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
 www-bio.lnl.gov/db/ftp/image/image.html

FEATURES  
 source  
 1..146  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="1256199"  
 /clone\_lib="NCI-CGAP-GC3"  
 /tissue\_type="pooled germ cell tumors"  
 /lab\_host="DH10B"  
 /note="vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker: 1st strand cDNA was prepared from 3 pooled  
 germ cell tumors, and was then primed with a Not I -  
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
 RI adaptors (Pharmacia), digested with Not I and cloned  
 into the Not I and Eco RI sites of the modified pT73  
 vector. Library is not normalized. Library was  
 constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 32 a 26 c 54 g 34 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gctggcgaggttggggcct 21  
 ||||||||||||||||  
 Db 106 gctggcgaggttggggcct 126

RESULT 7  
 LOCUS AA434384 151 bp mRNA linear EST 29-MAY-1997  
 DEFINITION zw31h1.s1 Soares ovary tumor NbHOT Homo sapiens CDNA clone  
 IMAGE:770949.3' similar to gb:U4723 CLUSTERIN PRECURSOR (HUMAN);,  
 mRNA sequence.  
 AA434384  
 AA434384.1 GI:2139298  
 EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 151)  
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
 Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie,  
 T., Waterston, R. and Wilson, R.

TITLE Washu-Merck EST Project 1997  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Wilton RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu  
 This clone is available royalty-free through LINL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Seq primer: -40m13 fwd. Et from Amersham.  
 Location/Qualifiers  
 1..151  
 /organism="Homo sapiens"  
 /db\_xref="GDB:5980822"  
 /db\_xref="taxon:9606"  
 /clone\_image="770949"  
 /clone\_lib="Soares ovary tumor NbHOT"  
 /sex="Female"  
 /tissue\_type="ovarian tumor"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Organ: ovary; Vector: pT73D (Pharmacia) with a  
 modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTTACCAATCTGAAGTGGAGCGCGCGGTTTTTTTTTTTTTTT 3']  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adaptors (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). Library constructed by Bento Soares and  
 M. Fatima Bonaldo. "

BASE COUNT 32 a 27 c 56 g 36 t  
 ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gctggcgaggttggggcct 21  
 ||||||||||||||||  
 Db 107 gctggcgaggttggggcct 127

RESULT 8  
 LOCUS BI055204 151 bp mRNA linear EST 15-JUN-2001  
 DEFINITION PM2-GN0495-050201-001-g05 GN0495 Homo sapiens CDNA, mRNA sequence.  
 ACCESSION BI055204  
 VERSION BI055204.1 GI:14462734  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 151)  
 AUTHORS Dias, Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,  
 Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,  
 Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H.,  
 Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare,  
 M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and  
 Simpson, A. U.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 CONTACT: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-27070422  
 Fax: +55-11-27070001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPSP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?rtl=PM2&cl2=PM2-GN0495-050201-001-g05&cl2=2001-02-05&cl4=1)

Seq primer: puc 18 forward  
 High quality sequence start: 7  
 High quality sequence stop: 148.

## FEATURES

Location/Qualifiers

1..151  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="GN0495"  
 /dev\_stage="Adult"

/note="Organ: Placenta,normal; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## BASE COUNT

24 a 32 c 61 g 34 t

## ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctggggggagttggggcct 21  
 |||  
 Db 61 GCTGGGGGAGTTGGGGCCT 81

## RESULT 9

BI040715 153 bp mRNA linear EST 14-JUN-2001  
 LOCUS PM2-WT0232-120201-005-d02 NT0232 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION  
 ACCESSION BI040715 GI:14447341  
 VERSION  
 KEYWORDS EST.

## SOURCE

human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 153)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brines,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,E.F., Goldman,G.R., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed

## TITLE

sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## JOURNAL

2020263

## MEDLINE

Contact: Simpson A.J.G.

## COMMENT

Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPSP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?rtl=PM2&cl2=PM2-WT0232-120201-005-d02&cl2=2001-02-12&cl4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 34  
 High quality sequence stop: 153.

## FEATURES

Location/Qualifiers

1..153  
 /organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="NT0232"

/dev\_stage="Adult"

/note="Organ: nervous,tumor; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## BASE COUNT

17 a 36 c 67 g 33 t

## ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 153;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctggggggagttggggcct 21  
 |||  
 Db 35 GCTGGGGGAGTTGGGGCCT 55

## RESULT 10

BI467181 157 bp mRNA linear EST 22-AUG-2001  
 LOCUS ic19a12.x4 HR85 islet Homo sapiens cDNA 3', mRNA sequence.  
 DEFINITION  
 ACCESSION BI467181  
 VERSION  
 KEYWORDS EST.

## SOURCE

human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 157)

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Maira,M., Page,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Rilter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)

## TITLE

Other-ESTs: ic19a12.y3

## JOURNAL

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557

Email: dmelton@bldhp.harvard.edu

Llibrary was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
 Washington University Genome Sequencing Center for information on  
 obtaining a clone please contact: Dr. Hiroshi Inoue  
 (hinoue@lm.wustl.edu)

## FEATURES

High quality sequence stop: 142.

## SOURCE

Location/Qualifiers

1..157

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="HR85 islet"

/tissue\_type="Purified pancreatic islet"

/lab\_host="DH10B"

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
 NotI; Site\_2: XhoI; cDNA made by oligo-dT priming.  
 Size-selected on agarose gel. Average insert size ~1kb. 5'

XhoI site was destroyed after directional cloning.  
 Amplified once. Contact information: Hiroshi Inoue, MD,  
 Metabolism Div. (Alan Permutt Lab), Washington University  
 School of Medicine, Box 8127, 660 South Euclid Ave., St.  
 Louis, MO 63110. E-mail: hinoue@ingate.wustl.edu, Tel:  
 314-362-1916, Fax: 314-747-2692."

```

BASE COUNT      35 a      28 c      56 g      38 t
ORIGIN

Query Match      100.0%; Score 21: DB 10; Length 157;
Best Local Similarity 100.0%; Pred. NO. 2.2e+02;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gctggcgaggttggggcct 21
|||||
Db 104 GCTGGCGGAGTTGGGGCCT 124

RESULT 11
LOCUS      BI040516
DEFINITION PM2-NT0232-070201-004-c03 NT0232 Homo sapiens CDNA, mRNA sequence.
ACCESSION  BI040516
VERSION     BI040516.1 GI:14447142
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
            1 (bases 1 to 159)
REFERENCE   1
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL    2002063
MEDLINE
COMMENT    Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM2&ct2=PM2-NT0232-
            070201-004-c03&v3=2001-02-07&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 15.
            Location/Qualifiers
                source
                    1..159
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone_id="NT0232"
                        /dev_stage="Adult"
                        /note="Organ: nervous tumor; Vector: puc18; Site 1: Smar;
                        derived from ORESTES PCR (U.S. Letters Patent application
                        No. 196,716 - Ludwig Institute for Cancer Research)
                        profiles into the puc 18 vector. Reverse transcription of
                        tissue mRNA and cdna amplification were performed under
                        low stringency conditions."
BASE COUNT      31 a      61 c      37 g      29 t      1 others
ORIGIN

Query Match      100.0%; Score 21: DB 10; Length 159;
Best Local Similarity 100.0%; Pred. NO. 2.2e+02;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gctggcgaggttggggcct 21
|||||

```

```

Db 134 GCTGGCGGAGTTGGGGCCT 114

RESULT 12
LOCUS      AI523434
DEFINITION ar72h08.x1 Barstead aorta HPLRB6 Homo sapiens CDNA clone
            IMAGE:2128191 3' similar to gb:xl4723 CLUSTERIN PRECURSOR (HUMAN);,
            mRNA sequence.
ACCESSION  AI523434
VERSION     AI523434.1 GI:4437569
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
            1 (bases 1 to 169)
REFERENCE   1
AUTHORS    Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacey,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
            WashU-NCI human EST project
            Unpublished (1997)
            Contact: Wilson RK
            Washington University School of Medicine
            444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@wustl.wustl.edu
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Glbco.
            Location/Qualifiers
                source
                    1..169
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone_id="IMAGE:2128191"
                        /dev_stage="adult, age 64"
                        /sex="male"
                        /lab_host="DH10B (phage resistant)"
                        /note="Organ: aorta; Vector: pTR73D-pac (pharmacia) with a
                        modified polylinker; Site 1: EcoRI; Site 2: NotI; 1st
                        strand cDNA was primed with a Not I - oligo(dT) primer [5'
                        TGTTCAGATCTGAGTGGAGCGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT
                        3']; double-stranded cDNA was ligated to Eco RI adaptors
                        [5' AATTCGATCGAAGC 3' and 5' GTTGCATCGC 3'], digested
                        with Not I and cloned into the Not I and Eco RI sites of
                        the modified pTR73 vector. Library constructed by Bob
                        Barstead."
BASE COUNT      36 a      33 c      59 g      41 t
ORIGIN

Query Match      100.0%; Score 21: DB 9; Length 169;
Best Local Similarity 100.0%; Pred. NO. 2.2e+02;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 gctggcgaggttggggcct 21
|||||
Db 108 GCTGGCGGAGTTGGGGCCT 128

RESULT 13
LOCUS      AI691105
DEFINITION tx90g01.x1 NCL CGAP ufa Homo sapiens CDNA clone IMAGE:2276880 3'
            similar to gb:xl4723 CLUSTERIN PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION  AI691105
VERSION     AI691105.1 GI:4902407
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens

```

REFERENCE 1 (bases 1 to 169)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 CDNA library preparation: Life Technologies, Inc.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: [www.bio.lnml.gov/bbrp/image/image.html](http://www.bio.lnml.gov/bbrp/image/image.html)  
 Insert length: 869 Std Error: 0.00  
 Seq primer: -40up from Gibco.

FEATURES  
 source  
 1..169  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="IMAGE:2276880"  
 /clone\_lib="NCI-CGAP, U14"  
 /tissue\_type="serous papillary carcinoma, high grade, 2 pooled tumors"  
 /lab\_host="DH10B"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site\_1: SalI; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.48 kb. Life Technologies catalog #: 11542-016"

BASE COUNT 33 a 27 c 49 g 60 t

ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 169;  
 Best Local Similarity 100.0%; Pred. NO. 2.2e+02;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 gctggcgaggttgaggcct 21  
 ||||||||||||||||||||  
 Db 140 GCTGGCGGAGTGTGGGGCCT 160

RESULT 14  
 BF897129 171 bp mRNA linear EST 18-JAN-2001  
 LOCUS IL0-MT0215-281100-501-all MT0215 Homo sapiens CDNA, mRNA sequence.  
 DEFINITION BF897129  
 ACCESSION BF897129  
 VERSION BF897129.1 GI:12288588  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 171)  
 Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bata, G. S., Simpson, D. H., Brunstein, A., deOliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.  
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 CONTACT: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922

TITLE  
 JOURNAL  
 MEDLINE  
 COMMENT

REFERENCE 1 (bases 1 to 171)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 CDNA library preparation: Life Technologies, Inc.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLT at: [www.bio.lnml.gov/bbrp/image/image.html](http://www.bio.lnml.gov/bbrp/image/image.html)  
 Insert length: 869 Std Error: 0.00  
 Seq primer: -40up from Gibco.

FEATURES  
 source  
 1..171  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="MT0215"  
 /dev\_stage="Adult"  
 /note="Organ: marrow; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 /716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 32 a 43 c 62 g 34 t

ORIGIN

Query Match 100.0%; Score 21; DB 10; Length 171;  
 Best Local Similarity 100.0%; Pred. NO. 2.2e+02;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 gctggcgaggttgaggcct 21  
 ||||||||||||||||||||  
 Db 21 GCTGGCGGAGTGTGGGGCCT 41

RESULT 15  
 BF919288 171 bp mRNA linear EST 19-JAN-2001  
 LOCUS IL0-MT0171-301000-501-all MT0171 Homo sapiens CDNA, mRNA sequence.  
 DEFINITION BF919288  
 ACCESSION BF919288  
 VERSION BF919288.1 GI:12315176  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 171)  
 Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bata, G. S., Simpson, D. H., Brunstein, A., deOliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.  
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 CONTACT: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?IL0-MT0215-281100-501-all&ts=2000-10-30&tl=1>)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 171.

TITLE  
 JOURNAL  
 MEDLINE  
 COMMENT

FEATURES  
 source  
 1..171  
 /organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone_lib="NT0171"
/dev_stage="Adult"
/note="Organ: nervous tumor; Vector: puc18; Site_1: Sma1;
Site_2: Sma1; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      32 a      43 c      62 g      34 t
ORIGIN

```

```

Query Match      100.0%; Score 21; DB 10; Length 171;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 gctggggcggaattggggcct 21
   |||11111111111111111111
Db 21 GCTGGGCGGAGTGGGGCCT 41

```

Search completed: May 17, 2002, 15:25:40  
 Job time: 14517 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 17:58:45 : Search time 147.64 Seconds  
(Without alignments)  
66.583 Million cell updates/sec

Title: US-09-913-325-12

Perfect score: 21

Sequence: 1 gctggcgagtggtgggacct 21

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 424068 segs, 234053524 residues

Total number of hits satisfying chosen parameters: 848136

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_New:\*

- 1: /cgn2\_6/ptodata1/pna/PCT\_NEW\_COMB.seq:\*
- 2: /cgn2\_6/ptodata1/pna/US06\_NEW\_COMB.seq:\*
- 3: /cgn2\_6/ptodata1/pna/US07\_NEW\_COMB.seq:\*
- 4: /cgn2\_6/ptodata1/pna/US08\_NEW\_COMB.seq:\*
- 5: /cgn2\_6/ptodata1/pna/US09\_NEW\_COMB.seq:\*
- 6: /cgn2\_6/ptodata1/pna/US10\_NEW\_COMB.seq:\*
- 7: /cgn2\_6/ptodata1/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	21	100.0	US-09-053-375B-227	Sequence 227, App
2	21	100.0	US-60-365-384-115	Sequence 115, App
3	21	100.0	US-10-116-255-17	Sequence 17, Appl
4	16.4	78.1	US-10-125-086-9	Sequence 9, Appl
5	16.4	78.1	PCT-US02-06990-10	Sequence 10, Appl
6	16.4	78.1	PCT-US02-06990-10	Sequence 10, Appl
7	16.4	78.1	PCT-US02-08239-10	Sequence 10, Appl
8	16.4	78.1	PCT-US02-09135-10	Sequence 10, Appl
9	16.4	78.1	PCT-US02-09105-10	Sequence 10, Appl
10	16.4	78.1	PCT-US02-09188-10	Sequence 10, Appl
11	16.4	78.1	PCT-US02-09257-10	Sequence 10, Appl
12	16.4	78.1	PCT-US02-09370-10	Sequence 10, Appl
13	16.4	78.1	PCT-US02-09785-10	Sequence 10, Appl
14	16.4	78.1	PCT-US02-09922-10	Sequence 10, Appl
15	16.4	78.1	US-10-103-313-10	Sequence 10, Appl
16	16.4	78.1	US-10-102-627-10	Sequence 10, Appl
17	16.4	78.1	US-10-103-295-10	Sequence 10, Appl
18	16.4	78.1	US-10-105-299-10	Sequence 10, Appl
19	16.4	78.1	US-10-106-698-10	Sequence 10, Appl
20	16.4	78.1	US-10-102-806-846	Sequence 846, App
21	16.4	78.1	US-10-112-857-10	Sequence 10, Appl
22	16.4	78.1	US-10-115-123-10	Sequence 10, Appl
23	16.4	78.1	US-10-115-928-10	Sequence 10, Appl
24	16.4	78.1	US-10-116-016-10	Sequence 10, Appl
25	16.4	78.1	US-10-116-355-10	Sequence 10, Appl
26	16.4	78.1	US-10-004-860-10	Sequence 10, Appl

27	16.4	78.1	US-10-125-451-30	Sequence 30, Appl
28	16.4	78.1	US-10-125-540-10	Sequence 10, Appl
29	16.4	78.1	US-10-074-045-10	Sequence 10, Appl
30	16.4	78.1	PCT-US02-01109-5	Sequence 5, Appl
31	16.4	78.1	PCT-US02-06990-5	Sequence 5, Appl
32	16.4	78.1	PCT-US02-09239-5	Sequence 5, Appl
33	16.4	78.1	PCT-US02-09135-5	Sequence 5, Appl
34	16.4	78.1	PCT-US02-09105-5	Sequence 5, Appl
35	16.4	78.1	PCT-US02-09188-5	Sequence 5, Appl
36	16.4	78.1	PCT-US02-09257-5	Sequence 5, Appl
37	16.4	78.1	PCT-US02-09370-5	Sequence 5, Appl
38	16.4	78.1	PCT-US02-09785-5	Sequence 5, Appl
39	16.4	78.1	PCT-US02-09922-5	Sequence 5, Appl
40	16.4	78.1	US-10-103-313-5	Sequence 5, Appl
41	16.4	78.1	US-10-102-627-5	Sequence 5, Appl
42	16.4	78.1	US-10-103-295-5	Sequence 5, Appl
43	16.4	78.1	US-10-105-299-5	Sequence 5, Appl
44	16.4	78.1	US-10-106-698-5	Sequence 5, Appl
45	16.4	78.1	US-10-102-806-841	Sequence 841, App

#### ALIGNMENTS

RESULT 1  
US-09-053-375B-227/c  
; Sequence 227, Application US/09053375B  
; GENERAL INFORMATION:  
; APPLICANT: Chenchik, Alex  
; APPLICANT: Bibilashvili, Robert  
; TITLE OF INVENTION: Nucleic Acid Arrays  
; FILE REFERENCE: CLON-006  
; CURRENT APPLICATION NUMBER: US/09/053, 375B  
; NUMBER OF SEQ ID NOS: 1543  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 227  
; LENGTH: 1416  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-053-375B-227

Query Match 100.0%; Score 21; DB 5; Length 1416;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctggcgagtggtgggacct 21  
|||||  
Db 1309 GCTGGCGGCGAGTTGGCGCCT 1289

RESULT 2  
US-60-365-384-115/c  
; Sequence 115, Application US/60365384  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Wang, Jian-rui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Zhang, Jie  
; APPLICANT: Zhou, Ping  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Ghosh, Malabika  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Ma, Yundong  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Meng, Gezhi  
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

FILE REFERENCE: 814  
CURRENT APPLICATION NUMBER: US/60/365,384  
CURRENT FILING DATE: 2002-03-14  
NUMBER OF SEQ ID NOS: 666  
SOFTWARE: pf-FL-genes Version 6.0  
SEQ ID NO 115  
LENGTH: 1839  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (132)..(1316)  
US-60-365-384-115

Query Match 100.0%; Score 21; DB 7; Length 1839;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gctgggagagttgggggacct 21  
|||||  
Db 1453 GCTGGCGGAGTTGGGGGCGCT 1433

RESULT 3  
US-10-116-255-17/c  
Sequence 17, Application US/10116255  
GENERAL INFORMATION:  
APPLICANT: NI et al.  
TITLE OF INVENTION: Uncoupling Protein Polynucleotides, Polypeptides, and  
FILE REFERENCE: PTO09P1  
CURRENT APPLICATION NUMBER: US/10/116,255  
CURRENT FILING DATE: 2002-04-05  
PRIOR APPLICATION NUMBER: 09/685,897  
PRIOR FILING DATE: 2000-10-11  
PRIOR APPLICATION NUMBER: PCT/US00/09534  
PRIOR FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: 60/128,701  
PRIOR FILING DATE: 1999-04-09  
PRIOR APPLICATION NUMBER: 60/142,821  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: 60/149,448  
PRIOR FILING DATE: 1999-08-18  
PRIOR APPLICATION NUMBER: 60/164,751  
PRIOR FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 66  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 17  
LENGTH: 2876  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-116-255-17

Query Match 100.0%; Score 21; DB 6; Length 2876;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gctgggagagttgggggacct 21  
|||||  
Db 2734 GCTGGCGGAGTTGGGGGCGCT 2714

RESULT 4  
US-10-125-086-9/c  
Sequence 9, Application US/10125086  
GENERAL INFORMATION:  
APPLICANT: MONTELABRO, Ronald  
APPLICANT: CRAIGO, Jodi K.  
APPLICANT: ISSEL, Charles  
APPLICANT: PUFFER, Bridget  
APPLICANT: HENNESSY, Kristina J.

APPLICANT: BROWN, Karen K.  
TITLE OF INVENTION: EIAV P26 DELETION VACCINE AND DIAGNOSTIC  
FILE REFERENCE: MONTELABRO ET AL.  
CURRENT APPLICATION NUMBER: US/10/125,086  
CURRENT FILING DATE: 2002-04-18  
PRIOR APPLICATION NUMBER: US/09/659,026A  
PRIOR FILING DATE: 2001-06-11  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 226  
TYPE: DNA  
ORGANISM: IN VITROGEN  
US-10-125-086-9

Query Match 78.1%; Score 16.4; DB 6; Length 226;  
Best Local Similarity 94.4%; Pred. No. 2.4e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ctgggagagttgggggacct 19  
|||||  
Db 168 CTGGGCGGAGTTAGGGGC 151

RESULT 5  
PCT-US02-01109-10/c  
Sequence 10, Application PC/TUS0201109  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: 50 Human Secreted Proteins  
FILE REFERENCE: P2016PCT2  
CURRENT APPLICATION NUMBER: PCT/US02/01109  
CURRENT FILING DATE: 2002-01-17  
PRIOR APPLICATION NUMBER: US 60/262,066  
PRIOR FILING DATE: 2001-01-18  
NUMBER OF SEQ ID NOS: 206  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10  
LENGTH: 256  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
NAME/KEY: Protein\_Bind  
OTHER INFORMATION: Synthetic promoter for use in biological assays; Includes NF-K  
PCT-US02-01109-10

Query Match 78.1%; Score 16.4; DB 1; Length 256;  
Best Local Similarity 94.4%; Pred. No. 2.4e+02;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ctgggagagttgggggacct 19  
|||||  
Db 123 CTGGGCGGAGTTAGGGGC 106

RESULT 6  
PCT-US02-06990-10/c  
Sequence 10, Application PC/TUS0206990  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PF534PCT  
CURRENT APPLICATION NUMBER: PCT/US02/06990  
CURRENT FILING DATE: 2002-03-08  
PRIOR APPLICATION NUMBER: 60/274,214  
PRIOR FILING DATE: 2001-03-09  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10  
LENGTH: 256

```
;
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Protein_Bind
; OTHER INFORMATION: Synthetic promoter for use in biological assays; includes NF-KB
PCT-US02-06990-10
```

```
Query Match          78.1%; Score 16.4; DB 1; Length 256;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ctgggcgaggttggggc 19
    |||||
DB 123 CTGGCGGAGTTAGGGC 106
```

```
RESULT 7
PCT-US02-09239-10/c
; Sequence 10, Application PC/TUS0209239
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS953PCT
; CURRENT APPLICATION NUMBER: PCT/US02/09239
; PRIORITY FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 09/950,082
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 09/950,083
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 380
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 10
; LENGTH: 256
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Protein_Bind
; OTHER INFORMATION: Synthetic promoter for use in biological assays; includes NF-KB
PCT-US02-09239-10
```

```
Query Match          78.1%; Score 16.4; DB 1; Length 256;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ctgggcgaggttggggc 19
    |||||
DB 123 CTGGCGGAGTTAGGGC 106
```

```
RESULT 8
PCT-US02-09135-10/c
; Sequence 10, Application PC/TUS0209135
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS956PCT
; CURRENT APPLICATION NUMBER: PCT/US02/09135
; PRIORITY FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 09/950,082
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 09/950,083
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 491
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 10
```

```
;
; LENGTH: 256
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Protein_Bind
; OTHER INFORMATION: Synthetic promoter for use in biological assays; includes NF-K
PCT-US02-09135-10
```

```
Query Match          78.1%; Score 16.4; DB 1; Length 256;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ctgggcgaggttggggc 19
    |||||
DB 123 CTGGCGGAGTTAGGGC 106
```

```
RESULT 9
PCT-US02-09105-10/c
; Sequence 10, Application PC/TUS0209105
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS951PCT
; CURRENT APPLICATION NUMBER: PCT/US02/09105
; PRIORITY FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 09/950,082
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 09/950,083
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 779
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 10
; LENGTH: 256
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Protein_Bind
; OTHER INFORMATION: Synthetic promoter for use in biological assays; includes NF-K
PCT-US02-09105-10
```

```
Query Match          78.1%; Score 16.4; DB 1; Length 256;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ctgggcgaggttggggc 19
    |||||
DB 123 CTGGCGGAGTTAGGGC 106
```

```
RESULT 10
PCT-US02-09188-10/c
; Sequence 10, Application PC/TUS0209188
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS952PCT
; CURRENT APPLICATION NUMBER: PCT/US02/09188
; PRIORITY FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 09/950,082
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 09/950,083
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 1732
; SOFTWARE: Patentln Ver. 2.0
```

```

: SEQ ID NO 10
: LENGTH: 256
:
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: Protein_Bind
: OTHER INFORMATION: Synthetic promoter for use in biological assays; includes NF- $\kappa$ B
: OTHER INFORMATION: binding sites.
: CCT-0502-09186-10

```

Query Match	78.18;	Score 16.4;	DB 1;	Length 256;
Best Local Similarity	94.48;	Pred. No. 2.4e+02;		
Matches 17; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

```

Qy      2 ctggcgagltcgggc 19
          |||||
Db     123 CTGGCGGAGTtAGGGC 106

```

```

RESULT 11
PCT-US02-09257-10/c
: Sequence 10, Application PC/TUS0209257
: GENERAL INFORMATION:
: APPLICANT: Human Genome Sciences, Inc.
: TITLE OF INVENTION: Human Secreted Proteins
: FILE REFERENCE: PS957PCT
: CURRENT APPLICATION NUMBER: PCT/US02/09257
: CURRENT FILING DATE: 2002-03-26
: PRIOR APPLICATION NUMBER: US 60/278,650
: PRIOR FILING DATE: 2001-03-27
: PRIOR APPLICATION NUMBER: US 09/950,082
: PRIOR FILING DATE: 2001-09-12
: PRIOR APPLICATION NUMBER: US 09/950,083
: PRIOR FILING DATE: 2001-09-12
: NUMBER OF SEQ. ID NOS: 994
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 256
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: Protein Bind
: OTHER INFORMATION: Synthetic promoter for use in biological assays; includes NF-KB
PCT-US02-09257-10

```

```

Query Match      78.1%   Score 16.4; DB 1; Length 256;
Best Local Similarity 94.4%   Pred. NO. 2,4+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
-
Oy      2 ctgggcgaggtctggggc 19
          |||||
db      123 ctggcgcgaggttggggc 106

```

```

RESULT 12
PCT-US02-09370-10/c
Sequence 10, Application PC/TUS0209370
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS954PCT
CURRENT APPLICATION NUMBER: PCT/US02/09370
CURRENT FILING DATE: 2002-03-26
PRIORITY APPLICATION NUMBER: US 60/278,650
PRIORITY FILING DATE: 2001-03-27
PRIORITY APPLICATION NUMBER: US 09/950,082
PRIORITY FILING DATE: 2001-09-12
PRIORITY APPLICATION NUMBER: US 09/950,083
PRIORITY FILING DATE: 2001-09-12
NUMBER OF SEQ ID NOS: 1834

```

```

: SOFTWARE: Patentia Ver. 2.0
: SEQ ID NO 10
: LENGTH: 256
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: Protein-Bind
: OTHER INFORMATION: Synthetic promoter for use in biological assays; includes NF-K
: OTHER INFORMATION: binding sites.
PCT-US02-09370-10

```

Query Match	78.1%;	Score 16.4;	DB 1;	Length 256;
Best Local Similarity	94.4%;	Pred. No. 2.4e+02;		
Matches 17; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY      2 ctggcgagltgggac 19
          ||| ||| ||| ||| ||| |||
Db     123 CTGGCGGAGTTAGGGC 106
```

```

RESULT 13
PCT-US02-09785-10/c
: Sequence 10. Application PC/TUS0209785
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS905PCT
CURRENT APPLICATION NUMBER: PCT/US02/09785
PCT FILING DATE: 2002-03-19
PCT APPLICATION NUMBER: US 60/331,287
PCT FILING DATE: 2001-11-13
PCT APPLICATION NUMBER: US 60/306,171
PCT FILING DATE: 2001-07-19
PCT APPLICATION NUMBER: US 60/277,340
PCT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 1130
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 256
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: Protein.BIIND
OTHER INFORMATION: Synthetic promoter for use in biological assays; includes NF-κB binding sites.
PCT-US02-09785-10

```

	Query Match	78.1%	Score 16.4	DB 1	Length 256
	Best Local Similarity	94.4%	Pred. No. 2,4e+02		
Matches	17	Conservative	0	Mismatches 1	Indels 0
QY	2 ctgagcgagagttgagagc	19	.		
Db	123 CTGGCGGAGGTTAGGGGC	106			

```

RESULT 14
PCT-US02-09922-10/c
; Sequence 10, Application PC/TUS0209922
;
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS95PCT
;
; CURRENT APPLICATION NUMBER: PCT/US02/09922
;
; CURRENT FILING DATE: 2002-03-26
;
; PRIOR APPLICATION NUMBER: US 60/278,650
;
; PRIOR FILING DATE: 2001-03-27
;
; PRIOR APPLICATION NUMBER: US 09/950,082
;
; PRIOR FILING DATE: 2001-09-12
;
; PRIOR APPLICATION NUMBER: US 09/950,083
;
; PRIOR FILING DATE: 2001-09-12

```

```

: NUMBER OF SEQ ID NOS: 1117
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 256
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: NAME/KEY: Protein_Bind
: OTHER INFORMATION: Synthetic promoter for use in biological assays; includes NF-KB
: OTHER INFORMATION: binding sites.
PCT-US02-09922-10

```

```

Query Match          78.1%; Score 16.4; DB 1; Length 256;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 ctgggcgagcttgaggc 19
   |||1111111111111111
Db 123 CTGGCGGAGTTAGGGCC 106

```

```

RESULT 15
US-10-103-313-10/C
: Sequence 10, Application US/10103313
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P0207C1
: CURRENT APPLICATION NUMBER: US/10/103,313
: CURRENT FILING DATE: 2002-03-12
: NUMBER OF SEQ ID NOS: 653
: Prior Application removed - See File Wrapper or Palm
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 256
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Synthetic promoter for use in biological assays; includes NF-KB
: OTHER INFORMATION: binding sites.
US-10-103-313-10

```

```

Query Match          78.1%; Score 16.4; DB 6; Length 256;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 ctgggcgagcttgaggc 19
   |||1111111111111111
Db 123 CTGGCGGAGTTAGGGCC 106

```

```

Search completed: May 17, 2002, 17:58:46
Job time: 9180 sec

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 17:56:00 ; Search time 5965.48 Seconds  
(without alignments)  
76.160 Million cell updates/sec

Title: US-09-913-325-12

Perfect score: 21

Sequence: 1 gctggcgcggagcttgggacct 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : Pending\_Patents\_MA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/US0905A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US0906A.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US0907A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US0908A.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US0909A.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US0910A.COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US0911A.COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US0912A.COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US0913A.COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US0914A.COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US0915A.COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/US0916A.COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/US0917A.COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/US0918A.COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/US0919A.COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/US0920A.COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/US0921A.COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/US0922A.COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/US0923A.COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/US0924A.COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/US0925A.COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/US0926A.COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/US0927A.COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/US0928A.COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/US0929A.COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/US0930A.COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/US0931A.COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/US0932A.COMB.seq.\*  
29: /cgn2\_6/ptodata/2/pna/US0933A.COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/US0934A.COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/US0935A.COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/US0936A.COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/US0937A.COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/US0938A.COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/US0939A.COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/US0940A.COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/US0941A.COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/US0942A.COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/US0943A.COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/US0944A.COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/US0945A.COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/US0946A.COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/US0947A.COMB.seq.\*

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	21	1 PCT-US000-04875-12	Sequence 12, Appl
2	21	100.0	21	US-09-913-325-12	Sequence 12, Appl
3	21	100.0	21	US-09-944-326-12	Sequence 12, Appl
4	21	100.0	21	US-10-080-794-12	Sequence 9475, Ap
5	21	100.0	166	US-09-534-840-10418	Sequence 10418, A
6	21	100.0	166	US-09-534-840-10418	Sequence 10418, A
7	21	100.0	180	US-09-534-840-10418	Sequence 10418, A
8	21	100.0	228	US-09-534-840-10418	Sequence 10418, A
9	21	100.0	236	US-09-534-840-10418	Sequence 10418, A
10	21	100.0	240	US-09-534-840-10418	Sequence 10418, A
11	21	100.0	240	US-09-534-840-10418	Sequence 10418, A
12	21	100.0	244	US-09-534-840-10418	Sequence 10418, A
13	21	100.0	244	US-09-534-840-10418	Sequence 10418, A
14	21	100.0	260	US-09-534-840-10418	Sequence 10418, A
15	21	100.0	261	US-09-534-840-10418	Sequence 10418, A
16	21	100.0	284	US-09-534-840-10418	Sequence 10418, A
17	21	100.0	286	US-09-534-840-10418	Sequence 10418, A
18	21	100.0	286	US-09-534-840-10418	Sequence 10418, A
19	21	100.0	289	US-09-534-840-10418	Sequence 10418, A
20	21	100.0	302	US-09-534-840-10418	Sequence 10418, A
21	21	100.0	302	US-09-534-840-10418	Sequence 10418, A
22	21	100.0	306	US-09-534-840-10418	Sequence 10418, A
23	21	100.0	308	US-09-534-840-10418	Sequence 10418, A
24	21	100.0	313	US-09-534-840-10418	Sequence 10418, A
25	21	100.0	327	US-09-534-840-10418	Sequence 10418, A
26	21	100.0	330	US-09-534-840-10418	Sequence 10418, A
27	21	100.0	350	US-09-534-840-10418	Sequence 10418, A
28	21	100.0	362	US-09-534-840-10418	Sequence 10418, A
29	21	100.0	362	US-09-534-840-10418	Sequence 10418, A
30	21	100.0	362	US-09-534-840-10418	Sequence 10418, A
31	21	100.0	362	US-09-534-840-10418	Sequence 10418, A

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

```
c 32      21      100.0      375      20      US-09-534-840-10443      Sequence 10443, A
c 33      21      100.0      377      25      US-09-652-125A-667      Sequence 667, App
c 34      21      100.0      384      16      US-09-248-797-2583      Sequence 2583, Ap
c 35      21      100.0      384      34      US-09-925-564-2583      Sequence 2583, Ap
c 36      21      100.0      388      17      US-09-332-782-8373      Sequence 8373, Ap
c 37      21      100.0      388      19      US-09-513-694-8373      Sequence 8373, Ap
c 38      21      100.0      391      31      US-09-821-710-1224      Sequence 1224, Ap
c 39      21      100.0      395      32      US-09-834-366-48205      Sequence 48205, A
c 40      21      100.0      395      58      US-60-197-873-48205      Sequence 48205, A
c 41      21      100.0      396      56      US-60-172-373-19074      Sequence 19074, A
c 42      21      100.0      397      16      US-09-289-768-5610      Sequence 5610, Ap
c 43      21      100.0      397      35      US-09-939-397-5610      Sequence 5610, Ap
c 44      21      100.0      398      16      US-09-248-797-5333      Sequence 5333, Ap
c 45      21      100.0      398      17      US-09-346-956-1857      Sequence 1857, Ap
```

## ALIGNMENTS

```
RESULT 1
PCT-US00-04875-12
```

```
; Sequence 12, Application PC/TUS0004875
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Gleave, Martin
```

```
; APPLICANT: Rennie, Paul S.
```

```
; APPLICANT: Miyake, Hideaki
```

```
; APPLICANT: Nelson, Colleen
```

```
; TITLE OF INVENTION: ANTISENSE THERAPY FOR CANCER
```

```
; FILE REFERENCE: UBC.P-020-WO
```

```
; CURRENT APPLICATION NUMBER: PCT/US00/04875
```

```
; CURRENT FILING DATE: 2000-02-25
```

```
; EARLIER APPLICATION NUMBER: 60/121,726
```

```
; EARLIER FILING DATE: 1999-02-26
```

```
; NUMBER OF SEQ ID NOS: 14
```

```
; SOFTWARE: Patentln Ver. 2.1
```

```
; SEQ ID NO 12
```

```
; LENGTH: 21
```

```
; TYPE: DNA
```

```
; ORGANISM: HUMAN
```

```
; FEATURE:
```

```
; OTHER INFORMATION: antisense TRPM-2 ODN
```

```
PCT-US00-04875-12
```

```
Query Match      100.0%; Score 21; DB 1; Length 21;
```

```
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
```

```
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1      1 gctggcgagagttgggggcct 21
```

```
Db      1      1 gctggcgagagttgggggcct 21
```

```
RESULT 2
US-09-913-325-12
```

```
; Sequence 12, Application US/09913325
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Gleave, Martin
```

```
; APPLICANT: Rennie, Paul S.
```

```
; APPLICANT: Miyake, Hideaki
```

```
; APPLICANT: Nelson, Colleen
```

```
; TITLE OF INVENTION: TRPM-2 ANTISENSE THERAPY
```

```
; FILE REFERENCE: UBC.P-020-WO
```

```
; CURRENT APPLICATION NUMBER: US/09/913,325
```

```
; CURRENT FILING DATE: 2001-10-29
```

```
; PRIOR APPLICATION NUMBER: 60/121,726
```

```
; PRIOR FILING DATE: 1999-02-26
```

```
; NUMBER OF SEQ ID NOS: 14
```

```
; SOFTWARE: Patentln Ver. 2.1
```

```
; SEQ ID NO 12
```

```
; LENGTH: 21
```

```
; TYPE: DNA
```

```
; ORGANISM: HUMAN
```

```
; FEATURE:
; OTHER INFORMATION: antisense TRPM-2 ODN
US-09-913-325-12
```

```
Query Match      100.0%; Score 21; DB 34; Length 21;
```

```
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
```

```
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1      1 gctggcgagagttgggggcct 21
```

```
Db      1      1 gctggcgagagttgggggcct 21
```

```
RESULT 3
US-09-944-326-12
```

```
; Sequence 12, Application US/09944326
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Gleave, Martin
```

```
; APPLICANT: Rennie, Paul S.
```

```
; APPLICANT: Miyake, Hideaki
```

```
; APPLICANT: Nelson, Colleen
```

```
; TITLE OF INVENTION: TRPM-2 ANTISENSE THERAPY
```

```
; FILE REFERENCE: UBC.P-020-2
```

```
; CURRENT APPLICATION NUMBER: US/09/944,326
```

```
; CURRENT FILING DATE: 2001-08-30
```

```
; PRIOR APPLICATION NUMBER: 60/121,726
```

```
; PRIOR FILING DATE: 1999-02-26
```

```
; PRIOR APPLICATION NUMBER: 09/913,325
```

```
; PRIOR FILING DATE: 2001-08-10
```

```
; NUMBER OF SEQ ID NOS: 14
```

```
; SOFTWARE: Patentln Ver. 2.1
```

```
; SEQ ID NO 12
```

```
; LENGTH: 21
```

```
; TYPE: DNA
```

```
; ORGANISM: HUMAN
```

```
; FEATURE:
```

```
; OTHER INFORMATION: antisense TRPM-2 ODN
```

```
US-09-944-326-12
```

```
Query Match      100.0%; Score 21; DB 35; Length 21;
```

```
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
```

```
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1      1 gctggcgagagttgggggcct 21
```

```
Db      1      1 gctggcgagagttgggggcct 21
```

```
RESULT 4
US-10-080-794-12
```

```
; Sequence 12, Application US/10080794
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Gleave, Martin
```

```
; APPLICANT: Rennie, Paul S.
```

```
; APPLICANT: Miyake, Hideaki
```

```
; APPLICANT: Nelson, Colleen
```

```
; TITLE OF INVENTION: TRPM-2 ANTISENSE THERAPY USING AN OLIGONUCLEOTIDE
```

```
; FILE REFERENCE: HAVING 2'-O-(2-METHOXY)ETHYL MODIFICATIONS
```

```
; CURRENT APPLICATION NUMBER: US/10/080,794
```

```
; CURRENT FILING DATE: 2002-02-22
```

```
; PRIOR APPLICATION NUMBER: 60/121,726
```

```
; PRIOR FILING DATE: 1999-02-26
```

```
; PRIOR APPLICATION NUMBER: 09/913,325
```

```
; PRIOR FILING DATE: 2001-08-10
```

```
; PRIOR APPLICATION NUMBER: 09/944,326
```

```
; PRIOR FILING DATE: 2001-08-30
```

```
; NUMBER OF SEQ ID NOS: 19
```

```
; SOFTWARE: Patentln Ver. 2.1
```

```
; SEQ ID NO 12
```



LENGTH: 21  
TYPE: DNA  
ORGANISM: HUMAN  
FEATURE:  
OTHER INFORMATION: antisense TRPM-2 ODN  
US-10-080-794-12

Query Match 100.0%; Score 21; DB 37; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctggcgagagttggggcct 21  
|||||

DB 1 gctggcgagagttggggcct 21

RESULT 5  
US-09-534-840-9475/c  
Sequence 9475, Application US/09534840  
GENERAL INFORMATION:  
APPLICANT: Sellhame, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MOLEC  
FILE REFERENCE: PD-1010 CIP  
CURRENT APPLICATION NUMBER: US/09/534,840  
CURRENT FILING DATE: 1992-07-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 13675  
SOFTWARE: PERL Program  
SEQ ID NO 9475  
LENGTH: 166  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No: hu00589179  
NAME/KEY: unsure  
LOCATION: 64, 73, 121, 140, 152, 160, 163  
OTHER INFORMATION: a, t, c, g, or other  
US-09-534-840-9475

Query Match 100.0%; Score 21; DB 20; Length 166;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctggcgagagttggggcct 21  
|||||

DB 27 gctggcgagagttggggcct 7

RESULT 6  
US-60-033-401-581/c  
Sequence 581, Application US/60033401  
GENERAL INFORMATION:  
APPLICANT: Gooding, Douglas H.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Ito, Laura Y.  
APPLICANT: Akerblom, Ingrid E.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Naughton, Rebecca E.  
APPLICANT: Klingler, Tod M.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
TITLE OF INVENTION: NEUROANGIOTIN TUMOR  
NUMBER OF SEQUENCES: 3552  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/60/033,401  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: CERRONE, MICHAEL C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PD-0297P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 855-0555  
TELEFAX: (415) 845-4166  
INFORMATION FOR SEQ ID NO: 581:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 166 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
CLONE: 2306438H1  
US-60-033-401-581

Query Match 100.0%; Score 21; DB 42; Length 166;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctggcgagagttggggcct 21  
|||||

DB 27 gctggcgagagttggggcct 7

RESULT 7  
US-09-534-840-10418/c  
Sequence 10418, Application US/09534840  
GENERAL INFORMATION:  
APPLICANT: Sellhame, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MO  
FILE REFERENCE: PD-1010 CIP  
CURRENT APPLICATION NUMBER: US/09/534,840  
CURRENT FILING DATE: 1992-07-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 13675  
SOFTWARE: PERL Program  
SEQ ID NO 10418  
LENGTH: 180  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No: hu01316725  
NAME/KEY: unsure  
LOCATION: 2-3  
OTHER INFORMATION: a, t, c, g, or other  
US-09-534-840-10418

Query Match 100.0%; Score 21; DB 20; Length 180;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 gctggcgagctggggcct 21  
|||||  
Db 73 GCTGGCGGAGTTGGGGCCT 53

RESULT 8  
US-09-534-840-10431/c  
; Sequence 10431, Application US/09534840  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MOLEC  
; FILE REFERENCE: PD-1010 CIP  
; CURRENT APPLICATION NUMBER: US/09/534,840  
; CURRENT FILING DATE: 1992-07-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 13675  
; SOFTWARE: PERL Program  
; SEQ ID NO 10431  
; LENGTH: 228  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc-feature  
; OTHER INFORMATION: Incyte ID No: hu01005556  
US-09-534-840-10431

Query Match 100.0%; Score 21; DB 20; Length 228;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 gctggcgagctggggcct 21  
|||||  
Db 126 GCTGGCGGAGTTGGGGCCT 106

RESULT 9  
US-09-534-840-10433/c  
; Sequence 10433, Application US/09534840  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MOLEC  
; FILE REFERENCE: PD-1010 CIP  
; CURRENT APPLICATION NUMBER: US/09/534,840  
; CURRENT FILING DATE: 1992-07-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 13675  
; SOFTWARE: PERL Program  
; SEQ ID NO 10433  
; LENGTH: 236  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc-feature  
; OTHER INFORMATION: Incyte ID No: hu01233159  
US-09-534-840-10433

Query Match 100.0%; Score 21; DB 20; Length 236;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 gctggcgagctggggcct 21  
|||||  
Db 134 GCTGGCGGAGTTGGGGCCT 114

RESULT 10  
US-09-534-840-10383/c  
; Sequence 10383, Application US/09534840  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MO  
; FILE REFERENCE: PD-1010 CIP  
; CURRENT APPLICATION NUMBER: US/09/534,840  
; CURRENT FILING DATE: 1992-07-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 13675  
; SOFTWARE: PERL Program  
; SEQ ID NO 10383  
; LENGTH: 240  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc-feature  
; OTHER INFORMATION: Incyte ID No: hu01273608  
US-09-534-840-10383

Query Match 100.0%; Score 21; DB 20; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 gctggcgagctggggcct 21  
|||||  
Db 132 GCTGGCGGAGTTGGGGCCT 112

RESULT 11  
US-09-534-840-10419/c  
; Sequence 10419, Application US/09534840  
; GENERAL INFORMATION:  
; APPLICANT: Seilhamer, Jeffrey J.  
; APPLICANT: Delegeane, Angelo M.  
; APPLICANT: Stuart, Susan G.  
; APPLICANT: Stuve, Laura L.  
; APPLICANT: Mullahy, Sara J.  
; APPLICANT: Naughton, Rebecca E.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING PROTEIN MODIFICATION MO  
; FILE REFERENCE: PD-1010 CIP  
; CURRENT APPLICATION NUMBER: US/09/534,840  
; CURRENT FILING DATE: 1992-07-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 13675  
; SOFTWARE: PERL Program  
; SEQ ID NO 10419  
; LENGTH: 240  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc-feature  
; OTHER INFORMATION: Incyte ID No: hu01048852  
US-09-534-840-10419

Query Match 100.0%; Score 21; DB 20; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctggcgagatgtggggcct 21  
 |||||  
 Db 132 GCTGGCGGAGTGTGGGGCCT 112

RESULT 12  
 US-09-534-840-10817/c  
 ; Sequence 10817, Application US/09534840  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sellhammer, Jeffrey J.  
 ; APPLICANT: Delegeane, Angelo M.  
 ; APPLICANT: Stuart, Susan G.  
 ; APPLICANT: Stuve, Laura L.  
 ; APPLICANT: Mullahy, Sara J.  
 ; APPLICANT: Naughton, Rebecca E.  
 ; FILE REFERENCE: PD-1010 CIP  
 ; CURRENT APPLICATION NUMBER: US/09/534,840  
 ; CURRENT FILING DATE: 1992-07-17  
 ; Prior application data removed - consult PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 13675  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 10817  
 ; LENGTH: 244  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No: hu01273243  
 US-09-534-840-10817

Query Match 100.0%; Score 21; DB 20; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctggcgagatgtggggcct 21  
 |||||  
 Db 134 GCTGGCGGAGTGTGGGGCCT 114

RESULT 13  
 US-09-534-840-10368/c  
 ; Sequence 10368, Application US/09534840  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sellhammer, Jeffrey J.  
 ; APPLICANT: Delegeane, Angelo M.  
 ; APPLICANT: Stuart, Susan G.  
 ; APPLICANT: Stuve, Laura L.  
 ; APPLICANT: Mullahy, Sara J.  
 ; APPLICANT: Naughton, Rebecca E.  
 ; FILE REFERENCE: PD-1010 CIP  
 ; CURRENT APPLICATION NUMBER: US/09/534,840  
 ; CURRENT FILING DATE: 1992-07-17  
 ; Prior application data removed - consult PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 13675  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 10368  
 ; LENGTH: 254  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No: hu01255852  
 US-09-534-840-10368

Query Match 100.0%; Score 21; DB 20; Length 254;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctggcgagatgtggggcct 21  
 |||||  
 Db 147 GCTGGCGGAGTGTGGGGCCT 127

RESULT 14  
 US-09-534-840-10355/c  
 ; Sequence 10355, Application US/09534840  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sellhammer, Jeffrey J.  
 ; APPLICANT: Delegeane, Angelo M.  
 ; APPLICANT: Stuart, Susan G.  
 ; APPLICANT: Stuve, Laura L.  
 ; APPLICANT: Mullahy, Sara J.  
 ; APPLICANT: Naughton, Rebecca E.  
 ; FILE REFERENCE: PD-1010 CIP  
 ; CURRENT APPLICATION NUMBER: US/09/534,840  
 ; CURRENT FILING DATE: 1992-07-17  
 ; Prior application data removed - consult PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 13675  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 10355  
 ; LENGTH: 260  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No: hu01203156  
 ; LOCATION: 26, 82, 92  
 ; OTHER INFORMATION: a, t, c, g, or other  
 US-09-534-840-10355

Query Match 100.0%; Score 21; DB 20; Length 260;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gctggcgagatgtggggcct 21  
 |||||  
 Db 183 GCTGGCGGAGTGTGGGGCCT 163

RESULT 15  
 US-09-652-125A-330/c  
 ; Sequence 330, Application US/09652125A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shyjan, Andrew W.  
 ; APPLICANT: Distefano, Peter  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
 ; FILE REFERENCE: 1600,1189-001  
 ; CURRENT APPLICATION NUMBER: US/09/652,125A  
 ; CURRENT FILING DATE: 2001-08-30  
 ; PRIOR APPLICATION NUMBER: 60/151,127  
 ; PRIOR FILING DATE: 1999-08-30  
 ; NUMBER OF SEQ ID NOS: 9506  
 ; SOFTWARE: FASTSEQ for windows Version 4.0  
 ; SEQ ID NO 330  
 ; LENGTH: 261  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(261)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-652-125A-330

Query Match 100.0%; Score 21; DB 25; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctgggcgagctgggggacct 21  
|||||  
Db 191 GCTGGCGGAGTGTGGGGGCT 171

Search completed: May 17, 2002, 17:56:01  
Job time: 12668 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:06:12 ; Search time 111.1 seconds  
(without alignments)

46,429 Million cell updates/sec

Title: US-09-913-325-12

Sequence: 1 gctggcgaggttgaggccct 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/pdata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/pdata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/pdata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/pdata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/pdata/2/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/pdata/2/ina/Backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	16.8	80.0	3562	4	US-09-360-197-1
C 2	16.8	80.0	3647	4	US-09-360-197-7
C 3	16.4	78.1	50	1	US-08-171-389-590
C 4	16.4	78.1	50	1	US-08-123-936-590
C 5	16.4	78.1	50	2	US-08-475-228A-590
C 6	16.4	78.1	50	3	US-08-482-080A-590
C 7	16.4	78.1	50	3	PCT-US93-12388-590
C 8	16.4	78.1	68	2	US-08-790-963-66
C 9	16.4	78.1	68	4	US-09-371-774-66
C 10	16.4	78.1	256	3	US-09-106-182-24
C 11	16.4	78.1	256	3	US-09-227-357-10
C 12	16.4	78.1	271	3	US-09-106-182-18
C 13	16.4	78.1	271	3	US-09-227-357-5
C 14	16.4	78.1	356	3	US-08-945-726-1
C 15	16.4	78.1	372	1	US-08-822-962-1
C 16	16.4	78.1	600	5	PCT-US91-02766-23
C 17	16.4	78.1	600	6	5219759-5
C 18	16.4	78.1	633	1	US-08-644-664B-6
C 19	16.4	78.1	633	4	US-08-761-277A-6
C 20	16.4	78.1	634	2	US-09-308-160B-25
C 21	16.4	78.1	664	1	US-08-455-755-1
C 22	16.4	78.1	664	2	US-07-910-260-1
C 23	16.4	78.1	664	5	PCT-US91-00878-1
C 24	16.4	78.1	2037	1	US-08-143-497-4
C 25	16.4	78.1	2037	1	US-08-461-666-4
C 26	16.4	78.1	2037	1	US-08-461-184-4
C 27	16.4	78.1	2037	1	US-08-463-675-4

#### ALIGNMENTS

C 28	16.4	78.1	2037	1	US-08-757-300-4	Sequence 4, Appl1
C 29	16.4	78.1	2037	1	US-08-464-589-4	Sequence 4, Appl1
C 30	16.4	78.1	2800	3	US-08-869-841B-42	Sequence 42, Appl1
C 31	16.4	78.1	3390	4	US-08-993-088A-5	Sequence 5, Appl1
C 32	16.4	78.1	3390	4	US-08-993-424B-5	Sequence 5, Appl1
C 33	16.4	78.1	3705	3	US-09-282-996-3	Sequence 3, Appl1
C 34	16.4	78.1	3944	1	US-07-678-408A-1	Sequence 1, Appl1
C 35	16.4	78.1	4059	2	US-08-252-493C-4	Sequence 4, Appl1
C 36	16.4	78.1	4059	3	US-09-276-197-4	Sequence 4, Appl1
C 37	16.4	78.1	4059	4	US-08-487-283A-3	Sequence 3, Appl1
C 38	16.4	78.1	4059	5	PCT-US95-07554-4	Sequence 4, Appl1
C 39	16.4	78.1	4059	5	PCT-US96-05611A-11	Sequence 11, Appl1
C 40	16.4	78.1	4454	1	US-07-712-284-1	Sequence 1, Appl1
C 41	16.4	78.1	4454	5	PCT-US92-04227-1	Sequence 1, Appl1
C 42	16.4	78.1	4665	3	US-08-948-378A-7	Sequence 7, Appl1
C 43	16.4	78.1	4665	4	US-09-169-425C-7	Sequence 7, Appl1
C 44	16.4	78.1	4696	2	US-08-929-967-15	Sequence 15, Appl1
C 45	16.4	78.1	4732	1	US-07-884-811-1	Sequence 1, Appl1

RESULT 1  
US-09-360-197-1/c  
Sequence 1, Application US/09360197  
Patent No. 6287859  
GENERAL INFORMATION:  
APPLICANT: Basillana, Frederic  
APPLICANT: Lazdunski, Michel  
APPLICANT: Waldmann, Rainer  
APPLICANT: Dewelle, Jan R.  
TITLE OF INVENTION: Human and Rat Families of Neuronal Acid-Sensitive  
FILE REFERENCE: 989, 6706P  
CURRENT FILING DATE: 1997-07-23  
PRIOR APPLICATION NUMBER: US/09/360,197  
PRIOR FILING DATE: 1998-08-05  
PRIOR APPLICATION NUMBER: 09/129,758  
PRIOR FILING DATE: 1998-08-05  
PRIOR APPLICATION NUMBER: 60/095,408  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 1  
LENGTH: 3562  
TYPE: DNA  
ORGANISM: rattus sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (123)..(1700)  
US-09-360-197-1  
Query Match 80.0%; Score 16.8; DB 4; Length 3562;  
Best local Similarity 90.0%; Pred. No. 65;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Db 1778 gctggcgaggttgaggcc 20  
Oy 1 gctggcgaggttgaggcc 20  
Db 1778 gctggcgaggttgaggcc 1759  
RESULT 2  
US-09-360-197-7/c  
Sequence 7, Application US/09360197  
Patent No. 6287859  
GENERAL INFORMATION:  
APPLICANT: Basillana, Frederic  
APPLICANT: Lazdunski, Michel  
APPLICANT: Waldmann, Rainer  
APPLICANT: Dewelle, Jan R.  
TITLE OF INVENTION: Human and Rat Families of Neuronal Acid-Sensitive  
FILE REFERENCE: 989, 6706P  
CURRENT FILING DATE: 1997-07-23  
PRIOR APPLICATION NUMBER: US/09/360,197  
PRIOR FILING DATE: 1998-08-05  
PRIOR APPLICATION NUMBER: 09/129,758  
PRIOR FILING DATE: 1998-08-05  
PRIOR APPLICATION NUMBER: 60/095,408  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 1  
LENGTH: 3562  
TYPE: DNA  
ORGANISM: rattus sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (123)..(1700)  
US-09-360-197-1

FILE REFERENCE: 989.6706P  
CURRENT APPLICATION NUMBER: US/09/360.197  
CURRENT FILING DATE: 1997-07-23  
PRIOR APPLICATION NUMBER: 09/129.758  
PRIOR FILING DATE: 1998-08-05  
PRIOR APPLICATION NUMBER: 60/095.408  
PRIOR FILING DATE: 1998-08-05  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 7  
LENGTH: 3647  
TYPE: DNA  
ORGANISM: ratcus sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (109)..(1785)  
US-09-360-197-7

Query Match 80.0% Score 16.8; DB 4; Length 3647;  
Best Local Similarity 90.0%; Pred. No. 65;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gctgggcgaggttggggccc 20  
||||| |  
Db 1863 GCTGGCGGCGAGTTGGCGGCC 1844

RESULT 3  
US-08-171-389-590/C  
Sequence 590, Application US/08171389  
Patent No. 5578444  
GENERAL INFORMATION:  
APPLICANT: Edwards, Cynthia A.  
APPLICANT: Cantor, Charles R.  
APPLICANT: Andrews, Beth M.  
APPLICANT: Turin, Lisa M.  
APPLICANT: Fry, Kirk E.  
TITLE OF INVENTION: Sequence-Directed DNA Binding  
TITLE OF INVENTION: Molecules, Compositions and Methods  
NUMBER OF SEQUENCES: 641  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/171.389  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/123.936  
FILING DATE: 17-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996.783  
FILING DATE: 23-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/723.618  
FILING DATE: 27-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/081.070  
FILING DATE: 22-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 590:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Simian virus 40 T/t late (start  
INDIVIDUAL ISOLATE: site 31)  
US-08-171-389-590

Query Match 78.1% Score 16.4; DB 1; Length 50;  
Best Local Similarity 94.4%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ctgggcgaggttggggccc 19  
||||| |  
Db 20 CTGGCGGCGAGTTAGCGGC 3

RESULT 4  
US-08-123-936-590/C  
Sequence 590, Application US/08123936  
Patent No. 5726014  
GENERAL INFORMATION:  
APPLICANT: Edwards, Cynthia A.  
APPLICANT: Cantor, Charles R.  
APPLICANT: Andrews, Beth M.  
APPLICANT: Turin, Lisa M.  
TITLE OF INVENTION: Screening Assay for the Detection of  
TITLE OF INVENTION: DNA-Binding Molecules  
NUMBER OF SEQUENCES: 640  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/123.936  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996.783  
FILING DATE: 23-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/723.618  
FILING DATE: 27-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 590:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE: NO  
INDIVIDUAL ISOLATE: Simian virus 40 T/t late (start  
INDIVIDUAL ISOLATE: site 31)  
US-08-123-936-590

Query Match 78.1%; Score 16.4; DB 1; Length 50;  
Best Local Similarity 94.4%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ctggcgagattgaggc 19  
|||||  
DB 20 CTGGCGGAGATTAGGGGC 3

RESULT 5  
US-08-475-228A-590/c  
Sequence 590, Application US/08475228A  
Patent No. 5869241  
GENERAL INFORMATION:  
APPLICANT: Edwards, Cynthia A.  
APPLICANT: Cantor, Charles R.  
APPLICANT: Andrews, Beth M.  
APPLICANT: Turin, Lisa M.  
APPLICANT: Fry, Kirk E.  
TITLE OF INVENTION: Sequence-Directed DNA Binding  
TITLE OF INVENTION: Molecules, Compositions and Methods  
NUMBER OF SEQUENCES: 664  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,228A  
FILING DATE: 06-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/123,936  
FILING DATE: 17-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996,783  
FILING DATE: 23-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/723,618  
FILING DATE: 27-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/081,070  
FILING DATE: 22-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stratford, Carol A.  
REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 590:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: Simian virus 40 T/t late (start  
INDIVIDUAL ISOLATE: site 31)  
US-08-475-228A-590

Query Match 78.1%; Score 16.4; DB 2; Length 50;  
Best Local Similarity 94.4%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ctggcgagattgaggc 19  
|||||  
DB 20 CTGGCGGAGATTAGGGGC 3

RESULT 6  
US-08-482-080A-590/c  
Sequence 590, Application US/08482080A  
Patent No. 6010849  
GENERAL INFORMATION:  
APPLICANT: Edwards, Cynthia A.  
APPLICANT: Cantor, Charles R.  
APPLICANT: Andrews, Beth M.  
APPLICANT: Turin, Lisa M.  
APPLICANT: Fry, Kirk E.  
TITLE OF INVENTION: Sequence-Directed DNA Binding  
TITLE OF INVENTION: Molecules, Compositions and Methods  
NUMBER OF SEQUENCES: 664  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,080A  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/171,389  
FILING DATE: 20-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/123,936  
FILING DATE: 17-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996,783  
FILING DATE: 23-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/723,618  
FILING DATE: 27-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/081,070  
FILING DATE: 22-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Brady, John F.  
REGISTRATION NUMBER: 39,118  
REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 324-0880  
TELEFAX: (650) 324-0960  
INFORMATION FOR SEQ ID NO: 590:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: Simian virus 40 T/t late (start)  
INDIVIDUAL ISOLATE: site 31)  
US-08-482-080A-590

Query Match 78.1%; Score 16.4; DB 3; Length 50;  
Best Local Similarity 94.4%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 2 ctggcgaggttggggc 19  
|||||  
Db 20 CTGGCGGAGTTAGGGCC 3

RESULT 7  
PCT-US93-12388-590/c  
Sequence 590, Application PC/TUS9312388  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Sequence-Directed DNA Binding  
TITLE OF INVENTION: Molecules, Compositions and Methods  
NUMBER OF SEQUENCES: 641  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genelabs Technologies, Inc.  
STREET: 505 Penobscot Drive  
CITY: Redwood City  
STATE: CA  
COUNTRY: USA  
ZIP: 94063  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/12388  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/123,936  
FILING DATE: 17-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/996,783  
FILING DATE: 23-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0860  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 590:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Simian virus 40 T/t late (start  
INDIVIDUAL ISOLATE: site 31)  
PCT-US93-12388-590

Query Match 78.1%; Score 16.4; DB 5; Length 50;  
Best Local Similarity 94.4%; Pred. No. 86;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 2 ctggcgaggttggggc 19  
|||||  
Db 20 CTGGCGGAGTTAGGGCC 3

RESULT 8  
US-08-790-963-66  
Sequence 66, Application US/08790963  
Patent No. 5837464  
GENERAL INFORMATION:  
APPLICANT: Daniel J. Capon  
ADDRESSEE: Christos John Petropoulos  
TITLE OF INVENTION: Compositions And Methods For  
TITLE OF INVENTION: Determining Anti-viral Drug Susceptibility And  
TITLE OF INVENTION: Resistance  
NUMBER OF SEQUENCES: 105  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/790,963  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 50130-B/JPW/ABC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0526  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-790-963-66

Query Match 78.1%; Score 16.4; DB 2; Length 68;  
Best Local Similarity 94.4%; Pred. No. 87;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 2 ctggcgaggttggggc 19  
|||||  
Db 33 CTGGCGGAGTTAGGGCC 50

RESULT 9  
US-09-371-774-66  
Sequence 66, Application US/09371774  
Patent No. 6242187  
GENERAL INFORMATION:  
APPLICANT: Daniel J. Capon  
ADDRESSEE: Christos John Petropoulos  
TITLE OF INVENTION: Compositions And Methods For  
TITLE OF INVENTION: Determining Anti-viral Drug Susceptibility And  
TITLE OF INVENTION: Resistance And Anti-viral Drug Screening  
NUMBER OF SEQUENCES: 105  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036  
COMPUTER READABLE FORM:



MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version#1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/371,774  
FILING DATE: 10-Aug-1999  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 50130-F/JPM/CMR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0526  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 68 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 66:  
US-09-371-774-66

Query Match 78.1%; Score 16.4; DB 4; Length 68;  
Best Local Similarity 94.4%; Pred. No. 87;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ctggcgagagttggggc 19  
|||||  
Db 33 CTGGCGGAGTGTAGGGC 50

RESULT 10  
US-09-106-182-24/c  
Sequence 24, Application US/09106182  
Patent No. 6046035  
GENERAL INFORMATION:  
APPLICANT: Shi, Yangu  
APPLICANT: Ruben, Steve  
TITLE OF INVENTION: Cardiotrophin-Like Cytokine  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc  
STREET: 9410 Key West Ave  
CITY: Rockville  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/106,182  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/051,053  
FILING DATE: 30-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET INFORMATION:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 256 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-106-182-24

Query Match 78.1%; Score 16.4; DB 3; Length 256;  
Best Local Similarity 94.4%; Pred. No. 90;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ctggcgagagttggggc 19  
|||||  
Db 123 CTGGCGGAGTGTAGGGC 106

RESULT 11  
US-09-227-357-10/c  
Sequence 10, Application US/09227357  
Patent No. 6342581  
GENERAL INFORMATION:  
APPLICANT: Fischer et al.  
TITLE OF INVENTION: 123 Human Secreted Proteins  
FILE REFERENCE: P2010P1  
CURRENT APPLICATION NUMBER: US/09/227,357  
EARLIER FILING DATE: 1999-01-08  
EARLIER APPLICATION NUMBER: PCT/US98/13684  
EARLIER FILING DATE: 1998-07-07  
EARLIER APPLICATION NUMBER: 60/051,926  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,793  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,925  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,929  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,803  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,732  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,931  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,932  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,916  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,930  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,918  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,920  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,733  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,795  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,919  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,928  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/055,722  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,723  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,948  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,949  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,953  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,950  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,947

EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,964  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/056,360  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,684  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,984  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,954  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/058,785  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058,664  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058,660  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058,661  
EARLIER FILING DATE: 1997-09-12  
NUMBER OF SEQ ID NOS: 672  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 10  
LENGTH: 256  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-227-357-10

Query Match 78.1% Score 16.4; DB 4; Length 256;  
Best Local Similarity 94.4%; Pred. No. 90;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 cctggcgagatctggggc 19  
|||||  
DB 123 CTGGCGAGATTAGGCGC 106

RESULT 12  
US-09-106-182-18/c  
Sequence 18, Application US/09106182  
Patent No. 6046035  
GENERAL INFORMATION:  
APPLICANT: Shi, Yanggu  
APPLICANT: Ruben, Steve  
TITLE OF INVENTION: Cardiotrophin-Like Cytokine  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc  
STREET: 9410 Key West Ave  
CITY: Rockville  
STATE: MD  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/106,182  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/051,053  
FILING DATE: 30-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PF385  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:  
LENGTH: 271 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-106-182-18

Query Match 78.1% Score 16.4; DB 3; Length 271;  
Best Local Similarity 94.4%; Pred. No. 90;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 cctggcgagatctggggc 19  
|||||  
DB 138 CTGGCGAGATTAGGCGC 121

RESULT 13  
US-09-227-357-5/c  
Sequence 5, Application US/09227357  
Patent No. 6342581  
GENERAL INFORMATION:  
APPLICANT: Fischer et al.  
TITLE OF INVENTION: 123 Human Secreted Proteins  
FILE REFERENCE: P2010P1  
CURRENT APPLICATION NUMBER: US/09/227,357  
CURRENT FILING DATE: 1999-01-08  
EARLIER APPLICATION NUMBER: PCT/US98/13684  
EARLIER FILING DATE: 1998-07-07  
EARLIER APPLICATION NUMBER: 60/051,926  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,793  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,925  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,929  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,803  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,732  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,931  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,932  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,916  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,930  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,918  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,920  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,733  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,795  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,919  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,928  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/055,722  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,723  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,948  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,949  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,953  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055,950

EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055, 947  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055, 964  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/056, 360  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055, 684  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055, 984  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/055, 954  
EARLIER FILING DATE: 1997-08-18  
EARLIER APPLICATION NUMBER: 60/058, 785  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058, 664  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058, 660  
EARLIER FILING DATE: 1997-09-12  
EARLIER APPLICATION NUMBER: 60/058, 661  
EARLIER FILING DATE: 1997-09-12  
NUMBER OF SEQ ID NOS: 672  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 5  
LENGTH: 271  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-227-357-5

Query Match 78.1%; Score 16.4; DB 4; Length 271;  
Best Local Similarity 94.4%; Pred. No. 90;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ctggcgagattgggggc 19  
|||||  
DB 138 CTGGCGGAGTTAGGGGC 121

RESULT 14  
US-08-945-726-1/C  
Sequence 1, Application US/08945726  
Patent No. 6004813  
GENERAL INFORMATION:  
APPLICANT: SERLUPPI-CRESCENZI, Ottaviano  
APPLICANT: PEZZOTTI, Annarita  
TITLE OF INVENTION: IL-6 ACTIVITY INHIBITORS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/945, 726  
FILING DATE: 10-NOV-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25, 618  
REFERENCE/DOCKET NUMBER: SERLUPPI-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 356 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-945-726-1

Query Match 78.1%; Score 16.4; DB 3; Length 356;  
Best Local Similarity 94.4%; Pred. No. 90;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ctggcgagattgggggc 19  
|||||  
DB 223 CTGGCGGAGTTAGGGGC 206

RESULT 15  
US-08-822-982-1  
Sequence 1, Application US/08822982  
Patent No. 5827705  
GENERAL INFORMATION:  
APPLICANT: Dean, David A.  
TITLE OF INVENTION: MOLECULE AND METHOD FOR IMPORTING DNA  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/822, 982  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/013, 863  
FILING DATE: 22-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/032, 468  
FILING DATE: 06-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BRAMAN, SUSAN J.  
REGISTRATION NUMBER: 34, 103  
REFERENCE/DOCKET NUMBER: 21467/32  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1600  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-822-982-1

Query Match 78.1%; Score 16.4; DB 1; Length 372;  
Best Local Similarity 94.4%; Pred. No. 90;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 2 ctggcgagattgggggc 19  
|||||

Mon May 20 10:17:49 2002

us-09-913-325-12.rni

Page 8

Db 134 CTGGCGGAGTTAGGGC 151

Search completed: May 17, 2002, 16:06:13  
Job time: 12747 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:13:55 ; Search time 447.74 Seconds  
(without alignments)  
80.527 Million cell updates/sec

Title: US-09-913-325-12

Perfect score: 21

Sequence: 1 gctggcgaggttggggcct 21

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : N.Geneseq\_032802.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	21	AAA94234	Human testostosterone
2	21	100.0	975	AAZ41353	Human normal uteru
3	21	100.0	1512	ABA83111	Apolipoprotein J o
4	21	100.0	1678	AAH23086	Osteoarthritis tis
5	21	100.0	2876	AAH25134	Human uncoupling p
6	16.8	80.0	2535	AAH25134	Nucleotide sequenc
7	16.8	80.0	2962	AAH28161	Rat acid sensitive
8	16.8	80.0	3562	AAV60839	Rat acid sensing i
9	16.8	80.0	3562	AAZ61197	CDNA encoding a ra

C 10	16.8	80.0	3647	19	AAV60842	Rat acid sensing i
C 11	16.8	80.0	3647	21	AAZ61200	CDNA encoding a ra
C 12	16.8	80.0	4121	21	AAZ75317	Human ORF ORF872
C 13	16.8	80.0	14654	22	AAZ82820	Genomic sequence #
C 14	16.4	78.1	50	15	AAO63840	Simian virus 40 T/
C 15	16.4	78.1	50	18	AAH64302	SV40 T/ Late (sta
C 16	16.4	78.1	50	20	AAH17590	Test sequence from
C 17	16.4	78.1	68	20	AAV64746	HIV anti-viral oli
C 18	16.4	78.1	68	22	AAH24758	Nucleotide sequenc
C 19	16.4	78.1	185	10	AAH90660	Sequence of SV40 e
C 20	16.4	78.1	204	21	AAH12021	Murine lactate deh
C 21	16.4	78.1	223	21	AAH12017	Human enolase A pr
C 22	16.4	78.1	231	21	AAH73859	SV40 promoter sequ
C 23	16.4	78.1	237	21	AAH12001	Murine PGK HRE der
C 24	16.4	78.1	242	20	AAZ07789	Promoter Obhrel us
C 25	16.4	78.1	242	21	AAH12016	Murine PGK HRE der
C 26	16.4	78.1	243	20	AAH13397	Murine PGK fragmen
C 27	16.4	78.1	243	21	AAH11995	Murine PGK HRE der
C 28	16.4	78.1	252	19	AAV52535	Sequence of SV40 e
C 29	16.4	78.1	256	19	AAV34153	Nuclear factor Kap
C 30	16.4	78.1	256	19	AAV59510	Nuclear factor Kap
C 31	16.4	78.1	256	19	AAV34285	Nuclear factor Kap
C 32	16.4	78.1	256	19	AAV69610	Nuclear factor Kap
C 33	16.4	78.1	256	20	AAZ19858	Nuclear factor KB-
C 34	16.4	78.1	256	20	AAZ24810	Chimeric NF-KB bin
C 35	16.4	78.1	256	20	AAZ09783	NF-kappaB/SV40 ear
C 36	16.4	78.1	256	20	AAZ10684	Nucleotide sequenc
C 37	16.4	78.1	256	20	AAZ00409	SV40 promoter temp
C 38	16.4	78.1	256	20	AAZ00801	SV40 early promote
C 39	16.4	78.1	256	20	AAZ06218	SV40 promoter cont
C 40	16.4	78.1	256	20	AAZ97915	Nuclear factor cont
C 41	16.4	78.1	256	20	AAZ97910	Nuclear factor Kap
C 42	16.4	78.1	256	20	AAZ82675	Nucleotide sequenc
C 43	16.4	78.1	256	20	AAZ84932	SV40 promoter cont
C 44	16.4	78.1	256	20	AAZ35900	NF-kappa B promote
C 45	16.4	78.1	256	20	AAZ37368	Human NF-kappaB pr

# ALIGNMENTS

RESULT 1	AAA94234	standard; DNA; 21 BP.
ID	AAA94234	
XX	AAA94234:	
AC	12-JAN-2001	(first entry)
DT	12-JAN-2001	
XX	Human testostosterone-repressed prostate message-2 antisense oligo #10.	
DE	Human: testostosterone-repressed prostate message-2; TRPM-2; clusterin;	
XX	KW sulfated glycoprotein-2; SGP-2; cancer; antisense oligonucleotide; ss.	
KW	Homo sapiens.	
OS	WO200049937-A2.	
PN	31-AUG-2000.	
PD	25-FEB-2000; 2000WO-US04875.	
XX	26-FEB-1999; 9905-0121726.	
XX	(UYBR-) UNIV BRITISH COLUMBIA.	
PA	Gleavage M, Rennie PS, Miyake H, Nelson C;	
XX	WPI; 2000-533132/48.	
DR	Treating prostatic tumors and renal cancers by antisense inhibition of	
XX	the testosterone-repressed prostate messenger-2 gene -	
PT		

PS Example 5: Page 38; 38pp; English.  
XX  
CC The present sequence is an antisense oligonucleotide directed at the  
CC human testosterone-repressed prostate message-2 (TRPM-2), also known as  
CC clusterin, sulfated glycoprotein-2 or SGP-2). The sequence was shown to  
CC promote the regression of tumours, and oligonucleotides directed  
CC at human TRPM-2 can be used in the treatment of tumour cells expressing  
CC the TRPM-2 gene. These include prostate cancer, renal cell cancer and  
CC some breast cancer cells. In addition to this, they also increase the  
CC chemosensitivity of the cells, meaning that conventional chemotherapy is  
CC more effective.  
XX  
SQ Sequence 21 BP; 1 A; 4 C; 12 G; 4 T; 0 other;  
  
Query Match 100.0%; Score 21; DB 21; Length 21;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 gctggcggaagtggggcct 21  
|||||  
DB 1 gctggcggaagtggggcct 21  
  
RESULT 2  
AAZ41353/C  
ID AAZ41353 standard; CDNA: 975 BP.  
XX  
AC AAZ41353;  
XX  
DT 19-JAN-2000 (first entry)  
XX  
DE Human normal uterus tissue derived CDNA 29.  
XX  
KW Human: uterus: cancer; treatment; anticancer; cytostatic; gene therapy;  
KM EST: expressed sequence tag; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
PN DE19817946-A1.  
XX  
PD 21-OCT-1999.  
XX  
PF 17-APR-1998; 98DE-1017946.  
XX  
PR 17-APR-1998; 98DE-1017946.  
XX  
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
DR WPI, 1999-591956/51.  
XX  
PT New nucleic acid sequences expressed in normal uterine tissues, and  
PT derived polypeptides, for treatment of uterine cancer and  
PT identification of therapeutic agents -  
XX  
PS Claim 3; Page 95; 154pp; German.  
XX  
CC This invention describes novel CDNA sequences (A) highly expressed in  
CC normal uterine tissue which can have anticancer and cytostatic activity  
CC and can be used for gene therapy. (A) are used (i) for recombinant  
CC expression of polypeptides (B) and (ii) to isolate complete genes.  
CC (B) are used (i) to identify agents suitable for treatment of uterine  
CC cancer; (ii) directly for treating this form of cancer (including of  
CC expression from gene therapy vectors) and (iii) for generation of  
CC specific antibodies. (A) are identified by assembling ESTs (expressed  
CC sequence tags) from a particular tissue type before comparison of  
CC expression patterns. This allows a significantly longer fragment of the  
CC gene to be revealed, so should reduce the number of failures associated  
CC with the fact that ESTs from different libraries may represent  
CC different parts of the same unknown gene, distorting the estimated  
CC frequency of occurrence in a particular tissue. AAZ41325-241385

CC represent the human uterine tissue derived CDNA fragments of the  
CC invention which encode the protein fragments represented in  
CC AAY9838-V59892.  
XX  
SQ Sequence 975 BP; 221 A; 320 C; 248 G; 186 T; 0 other;  
  
Query Match 100.0%; Score 21; DB 20; Length 975;  
Best Local Similarity 100.0%; Pred. No. 8.3;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 gctggcggaagtggggcct 21  
|||||  
DB 903 gctggcggaagtggggcct 883  
  
RESULT 3  
ABA83111/C  
ID ABA83111 standard; DNA: 1512 BP.  
XX  
AC ABA83111;  
XX  
DT 08-FEB-2002 (first entry)  
XX  
DE Apolipoprotein J ovarian tumour marker gene, SEQ ID NO:59.  
XX  
KW Ovarian tumour marker gene; human; overexpression; upregulation;  
KW epithelial tumour; cancer; diagnosis; prognosis; disease monitoring;  
KW identification; serous cystadenoma; borderline serous tumour;  
KW mucinous cystadenocarcinoma; mucinous cystadenocarcinoma;  
KW undifferentiated carcinoma; clear cell adenocarcinoma; cystadenofibroma;  
KW adenofibroma; Brenner tumour; serial analysis of gene expression; SAGE;  
KW immune response pathway; cell proliferation regulation; protein folding;  
KW membrane localized; secreted; therapeutic target; cytostatic;  
KW gene therapy; vaccine; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
PN WO200175177-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 03-APR-2001; 2001WO-US10947.  
XX  
PR 03-APR-2000; 2000US-194336P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PI Morin PJ, Sherman-Baust CA, Pizer ES, Hough CD;  
DR WPI: 2001-626450/72.  
XX  
PT Detecting and identifying ovarian tumor, identifying increased risk for  
PT developing ovarian cancer, and determining effectiveness of ovarian  
PT cancer treatment, by measuring expression level of ovarian tumor marker  
PT gene -  
XX  
PS Claim 23; Page 103; 140pp; English.  
XX  
CC The invention relates to methods for diagnosing and prognosing ovarian  
CC tumours in an individual via the detection and measurement of the  
CC expression of ovarian tumor marker genes (ABA83081-ABA83122, ABA83180,  
CC ABA83182 and ABA83184) or segments thereof (ABA83123-ABA83159, ABA83179,  
CC ABA83181 and ABA83183). The methods of the invention are useful for  
CC detecting an ovarian tumor in a patient, for identifying an individual  
CC at increased risk for developing ovarian cancer, in prognostic tests for  
CC assessing the relative severity of ovarian cancer, in tests for  
CC monitoring a patient in remission from ovarian cancer and in tests for  
CC monitoring disease status in a patient being treated for ovarian cancer.  
CC The methods can additionally be used to identify a particular tumour as  
CC being an ovarian tumour (i.e., an epithelial ovarian tumour selected from

CC serous cystadenoma, borderline serous tumour, serous cystadenocarcinoma,  
CC mucinous cystadenoma, borderline mucinous tumour, mucinous  
CC cystadenocarcinoma, endometrioid carcinoma, undifferentiated carcinoma,  
CC clear cell adenocarcinoma, cystadenofibroma, adenofibroma and Brenner  
CC tumour. The ovarian tumour marker genes of the invention were identified  
CC using SAGE (serial analysis of gene expression) and were found to be  
CC overexpressed in a broad variety of ovarian epithelial tumour cells  
CC relative to normal ovarian epithelial cells. The marker genes are  
CC implicated in immune response pathways, in the regulation of cell  
CC proliferation and in protein folding, and many of these are membrane-  
CC localised or secreted. In addition to their use as diagnostic and  
CC prognostic markers, the ovarian tumour marker genes or their encoded  
CC proteins may be used as therapeutic targets for the treatment and  
CC prevention of ovarian cancer. Sequences ABA83081-ABA83122, ABA83180,  
CC ABA83182 and ABA83184 represent the ovarian tumour marker genes of  
CC the invention.

SO Sequence 1512 BP; 378 A; 461 C; 387 G; 286 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 1512;

Best Local Similarity 100.0%; Pred. No. 8.2;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gctggcgaggttgaggcct 21

Db 1405 GCTGGCGAGTGTGGGCGCT 1385

RESULT 4  
AAH23086/c  
ID AAH23086 standard; DNA; 1678 BP.

XX  
XX AAH23086;

DT 17-SEP-2001 (first entry)

DE Osteoarthritis tissue-derived nucleic acid sequence #16.

KM Osteoarthritis; infectious disorder; Crohn's disease; sepsis; human;

KM wound healing; osteopathic; anti-arthritis; anti-inflammatory;

KM vulnery; antibacterial; antiallergic; ds.

OS Homo sapiens.

XX  
XX WO200153531-A2.

PD 26-JUL-2001.

XX  
XX 18-JAN-2001; 2001WO-US00016.

XX  
XX 18-JAN-2000; 2000US-0176523.

XX  
XX (PHAA ) PHARMACIA CORP.

XX  
XX Philpard D, Vasanthakamur G, Dotson S, Ma X;

DR WPI; 2001-451914/48.

PT Substantially purified protein, polypeptide or their fragments, used to  
PT identify a biologically active compound or composition and treat  
PT mammalian osteoarthritis -

PS Claim 1; Page 102-103; 144pp; English.

XX  
XX Sequences AAH23071-23152 represent nucleic acid sequences derived from  
CC osteoarthritis tissues. The sequences are useful as probes and for the  
CC diagnosis or prognosis of mammalian osteoarthritis. The polynucleotides  
CC and polypeptides of the invention are useful for generating diagnostic  
CC reagents, as targets for small molecule drug development, generation of  
CC therapeutics, and cloning genes. Specific antibodies are used to generate  
CC enzyme linked immunosorbent assays for detection of osteoarthritis. The  
CC invented molecules can be used to treat osteoarthritis or to analyse the

CC disease-modifying activity of osteoarthritis drugs. Other disorders  
CC treatable using the nucleic acid sequences include atopic, inflammatory  
CC and infectious disorders e.g. Crohn's disease and sepsis, and wound  
CC healing.

SO Sequence 1678 BP; 424 A; 491 C; 450 G; 313 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 1678;

Best Local Similarity 100.0%; Pred. No. 8.2;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gctggcgaggttgaggcct 21

Db 1558 GCTGGCGAGTGTGGGCGCT 1538

RESULT 5  
AAC90467/c

XX  
XX AAC90467 standard; cDNA; 2876 BP.

AC AAC90467;

DT 12-MAR-2001 (first entry)

DE Human uncoupling protein cDNA #16.

XX  
XX Human; uncoupling protein; immunosuppressive; antiarthritis;

KM antirheumatic; antiproliferative; cardiac; vasotropic;

KM cerebroprotective; neuroprotective; antibacterial; opthalmological;

KM gastrointestinal; nephrotoxic; gynaecological; vulnery; thrombolytic;

KM gene therapy; cancer; wound; infectious disease; thrombosis; arthritis;

KM infertility; ss.

XX  
XX Homo sapiens.

XX  
XX WO200061614-A2.

XX  
XX 19-OCT-2000.

XX  
XX 06-APR-2000; 2000WO-US09534.

XX  
XX 09-APR-1999; 99US-0128701.

XX  
XX 08-JUL-1999; 99US-0142821.

XX  
XX 18-AUG-1999; 99US-0149448.

XX  
XX 12-NOV-1999; 99US-0164751.

XX  
XX (HUMA-) HUMAN GENOME SCI INC.

XX  
XX Ruben SM, Ni J, Komatsoulis G, Rosen CA, Soppet DR;

XX  
XX WPI; 2000-656322/63.

XX  
XX P-PSDB; AAB50393.

XX  
XX Uncoupling proteins and nucleic acid sequences encoding them, useful  
PT for detecting, preventing and treating proliferative, neurological,  
PT immune system, cardiovascular and gastrointestinal disorders -

PS Claim 1; Page 313-314; 343pp; English.

XX  
XX The present sequence is one of eighteen isolated nucleotide sequences  
CC encoding uncoupling proteins. The nucleotide sequences may be used for  
CC the detection of various disorders such as cancer, for chromosome  
CC identification, as chromosome markers and for numerous other diagnostic  
CC or research purposes. The uncoupling protein encoded by the nucleotide  
CC sequences may be used to treat disorders such as neural, immune,  
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
CC renal and proliferative disorders, wounds, infectious diseases,  
CC thrombosis, arthritis, and infertility.

XX  
XX Sequence 2876 BP; 775 A; 695 C; 669 G; 737 T; 0 other;

Query Match 100.0%; Score 21; DB 21; Length 2876;  
 Best Local Similarity 100.0%; Pred. No. 8.1;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 gctggcgagctgggggacct 21  
 |||||||  
 Db 2734 GCTGGCGGAGCTTGGGGCGCT 2714

RESULT 6  
 AAH25134/C  
 ID AAH25134 standard; DNA: 2535 BP.  
 XX  
 AC AAH25134;  
 XX  
 DT 22-AUG-2001 (first entry)  
 XX  
 DE Nucleotide sequence of human alcohol dehydrogenase 21612.  
 XX  
 KM Human: alcohol dehydrogenase; colon disorder; brain disorder;  
 KM skin disorder; heart disorder; blood vessel disorder; kidney disorder;  
 KM prostate disorder; skeletal muscle disorder; ovary disorder;  
 KM testis disorder; epididymis disorder; spleen disorder; lung disorder;  
 KM liver disorder; uterus disorder; endometrium disorder; T-cell disorder;  
 KM red cell disorder; thymus disorder; B cell disorder; breast disorder;  
 KM thyroid disorder; pancreas disorder; small intestine disorder;  
 KM reduced platelet number disorder; precursor T cell neoplasm; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 762..2018  
 FT /\*tag= a  
 FT /\*product= "alcohol dehydrogenase 21612"  
 XX  
 PN W020014446-A2.  
 XX  
 PD 21-JUN-2001.  
 XX  
 PF 15-DEC-2000; 2000MO-US33873.  
 XX  
 PR 15-DEC-1999; 98US-0464039.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Meyers R;  
 XX  
 DR WPI: 2001-390244/41.  
 DR P-PSDB: AAB84367.  
 XX  
 PT Novel human alcohol dehydrogenase proteins; 21612, 21615, 21620, 21676,  
 PT 33756, useful for treating psoriasis; tropical sprue; pancreatitis,  
 PT golfer, osteomalacia, endometriosis, angina pectoris, embolism  
 XX  
 PS Claim 2; Fig 15; 156pp; English.  
 XX  
 CC AAH25131-35 encode human alcohol dehydrogenase proteins, designated  
 CC 21620, 33756, 21676, 21612 and 21615, respectively. Alcohol dehydrogenase  
 CC polynucleotides and polypeptides are useful for treatment and diagnosis  
 CC of disorders mediated by or related to alcohol dehydrogenase. They can  
 CC be used for treating disorders of colon, brain, skin, heart, blood  
 CC vessels, kidney, prostate, skeletal muscle, ovary, testis and epididymis,  
 CC spleen, lung, liver, uterus and endometrium, T-cells, red cells, thymus,  
 CC B cells, breast, thyroid, pancreas, small intestine, reduced platelet  
 CC number, precursor T cell neoplasms, bone forming cells, and bone marrow  
 CC cells.  
 XX  
 SQ Sequence 2535 BP; 720 A; 531 C; 537 G; 695 T; 52 other;

Query Match 80.0%; Score 16.8; DB 22; Length 2535;  
 Best Local Similarity 90.0%; Pred. No. 4.2e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 gctggcgagctgggggacct 20  
 |||||||  
 Db 73 GATGGCGGAGCTTGGGGCGCC 54

RESULT 7  
 AAH28161/C  
 ID AAH28161 standard; cDNA to mRNA; 2962 BP.  
 XX  
 AC AAH28161;  
 XX  
 DT 16-JUN-1999 (first entry)  
 XX  
 DE Rat Acid sensitive ion channel coding sequence.  
 XX  
 KM Acid sensitive ion channel; rat; ASIC; pH mediated pain disorder;  
 KM ischemia; gene therapy; proton-gated ion channel; ds.  
 XX  
 OS Rattus sp.  
 XX  
 PN W09911784-A1.  
 XX  
 PD 11-MAR-1999.  
 XX  
 PR 28-AUG-1998; 98MO-GB02609.  
 XX  
 PR 29-AUG-1997; 97GB-0018365.  
 XX  
 PA (UNIO ) UNIV COLLEGE LONDON.  
 XX  
 PI Akopian AN, Chen C, England S, Wood JN;  
 XX  
 DR WPI: 1999-205188/17.  
 DR P-PSDB: AAY03186.  
 XX  
 PT Acid sensitive ion channel (ASIC) proteins - useful in gene therapy  
 PT for treatment of pH mediated pain disorders  
 XX  
 PS Claim 7; Page 43-47; 62pp; English.  
 XX  
 CC This sequence encodes an acid sensitive ion channel (ASIC) of the  
 CC CC invention, isolated from rat dorsal root ganglion. Antisense ASIC DNA or  
 CC RNA is useful in gene therapy for downgrading expression of ASIC protein,  
 CC for pH mediated pain disorders e.g. in ischemia. The vectors are useful  
 CC for incorporating ASIC DNA or RNA for use in gene therapy. Proton-gated  
 CC ion channel agonists and antagonists are identified using cells  
 CC transformed with ASIC DNA by allowing interaction between the candidate  
 CC substance and ASIC protein in the membrane, and measuring interaction  
 CC and/or cell response. Partial agonists and antagonists can be identified  
 CC by their ability to block the response of the cell to present in a  
 CC solution of a given acid pH or any agonist. The hybridisation probes are  
 CC useful for screening libraries for ASIC DNA or RNA.  
 XX  
 SQ Sequence 2962 BP; 655 A; 850 C; 785 G; 672 T; 0 other;

Query Match 80.0%; Score 16.8; DB 20; Length 2962;  
 Best Local Similarity 90.0%; Pred. No. 4.2e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 gctggcgagctgggggacct 20  
 |||||||  
 Db 1824 GCTGGCGGAGCTTGGGGCGCC 1805

RESULT 8  
 AAV60839/C  
 ID AAV60839 standard; DNA; 3562 BP.  
 XX  
 AC AAV60839;  
 XX  
 DT 02-FEB-1999 (first entry)



XX Rat acid sensing ionic channel ASIC1A gene.

DE Rat; neuronal cationic channel; amiloride; proton; ASIC; brain; probe;  
XX acid sensing ionic channel; hybridisation; primer; PCR; amplification;  
KW modulator; acidity; nociception; pain; taste; inflammation; ischaemia;  
KW tumour; cerebral neurodegeneration; transgenic animal; knockout animal;  
KW gene therapy; Alzheimer's; Parkinson's; Huntington's; disease;  
XX amyotrophic lateral sclerosis; cerebellar ataxia; ds.

OS Rattus sp.

XX Key Location/Qualifiers  
FH 123..1703  
FT CDS /tag= a  
FT /product= "ASIC1A"  
FT /note= "acid sensing ionic channel 1A"

XX MO9835034-A1.

XX 13-AUG-1998.

XX 11-FEB-1998; 98WO-FR00270.

XX 28-JUL-1997; 97FR-0009587.

XX 11-FEB-1997; 97FR-0001574.

XX (CNRS ) CNRS CENT NAT RECH SCI.

XX Bassilana F, Champigny G, Heurteaux C, Lazdunski M;  
PI Waldmann R, Lingueglia E;

XX WPI: 1998-447231/38.

XX P-PSDB; AAM68504.

XX Protein comprising proton-sensitive neuronal channel - useful for  
PT screening for analgesics and for treating neurodegeneration

XX Claim 12; Page 24-27; 64pp; French.

XX This sequence represents the gene encoding a rat neuronal cationic  
CC channel that is sensitive to amiloride and activated by protons and  
CC is designated Acid Sensing Ionic Channel (ASIC). This sequence  
CC represents the rat ASIC1A gene. The sequence was isolated from a rat  
CC brain DNA library using a probe amplified by primers AAV60845-V60846.  
CC The protein can be used to screen for modulators of these channels,  
CC particularly to identify compounds that modulate perception of acidity,  
CC as regards nociception (pain) rather than taste. These compounds are  
CC used to treat or prevent pain associated with acidity (e.g. in cases of  
CC inflammation, ischaemia or some tumours) and as inhibitors of  
CC neurodegeneration caused by overexpression of the channels. Antibodies  
CC to the protein are used to detect the channels in tissues, and to act  
CC therapeutically as channel modulators. The nucleic acid can be used to  
CC generate transgenic, particularly knockout, animals for studying  
CC ASIC-related disorders, also for gene therapy. The channel protein,  
CC or its (anti)agonists, can be used to treat or prevent cerebral  
CC neurodegenerative conditions (e.g. Alzheimer's, Parkinson's or  
CC Huntington's diseases, amyotrophic lateral sclerosis or cerebellar  
CC ataxia.

XX Sequence 3562 BP; 905 A; 827 C; 862 G; 968 T; 0 other;

XX Query Match

XX Best Local Similarity 80.0%; Score 16.8; DB 19; Length 3562;  
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX 1 gctggcgagcttggggcc 20

DB 1778 GCTGGGCGAGCTTGGGGGCC 1759

RESULT 9

AAZ61197/c  
ID AAZ61197 standard; cDNA; 3562 BP.

XX AAZ61197;

XX 30-MAY-2000 (first entry)

XX cDNA encoding a rat acid-sensitive cationic channel 1A (rASIC1A).

XX Neuronal acid-sensitive cation channel; ASIC; ASIC 1A;

XX proton-gated cation channel; biphasic desensitisation; amiloride;  
KW cation transport channel; acid sensor; pH detection; ds.

XX Rattus sp.

XX Key Location/Qualifiers  
FH 123..1703  
FT CDS /tag= a  
FT /product= "acid-sensitive cationic channel 1A"

XX WO200008149-A2.

XX 17-FEB-2000.

XX 05-AUG-1999; 99WO-IB01445.

XX 05-AUG-1998; 98US-0095408.

XX (CNRS ) CNRS CENT NAT RECH SCI.

XX Waldmann R, Bassilana F, Lazdunski M, De Weille JR;

XX WPI: 2000-195574/17.

XX P-PSDB; AAV69175.

XX Novel human cation transport protein, Acid Sensing Ionic Channel 3 used  
PT to identify substances capable of modulating cation transport channel  
PT activity

XX Disclosure; Page 63-66; 84pp; English.

XX The present sequence encodes a rat neuronal acid-sensitive cation  
CC channel 1A (ASIC1A) protein. The protein is a proton-gated cation  
CC channel subunit that has biphasic desensitisation kinetics with both  
CC a rapidly inactivating sodium-selective and a sustained component. The  
CC channels are sensitive to amiloride. The specification describes ASIC3  
CC proteins, which are expressed in the sensory neurons but not in the  
CC brain. The cation transport channel proteins can be used in methods to  
CC identify substances capable of modulating the activity of cation  
CC transport channels. The human ASIC3 protein is also an acid sensor,  
CC and might play an important role in the detection of lasting pH changes  
CC in humans.

XX Sequence 3562 BP; 905 A; 827 C; 862 G; 968 T; 0 other;

XX Query Match

XX Best Local Similarity 80.0%; Score 16.8; DB 21; Length 3562;  
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

XX 1 gctggcgagcttggggcc 20

DB 1778 GCTGGGCGAGCTTGGGGGCC 1759

XX RESULT 10

XX AAV60842/c  
XX AAV60842 standard; DNA; 3647 BP.

XX AAV60842;

XX 02-FEB-1999 (first entry)

DE Rat acid sensing ionic channel ASIC1B gene.  
 XX  
 KW Rat: neuronal cationic channel; amiloride; proton; ASIC; brain; probe;  
 KW acid sensing ionic channel; hybridisation; primer; PCR; amplification;  
 KW modulator; acidity; nociception; pain; taste; inflammation; ischaemia;  
 KW tumour; cerebral neurodegeneration; transgenic animal; knockout animal;  
 KW gene therapy; Alzheimer's; Parkinson's; Huntington's; disease;  
 KW amyotrophic lateral sclerosis; cerebellar ataxia; ds.  
 XX  
 OS Rattus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 109..1788  
 FT /tag= a  
 FT /product= "ASIC1B"  
 FT /note= "acid sensing ionic channel 1B"  
 XX  
 PN W09835034-A1.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PE 11-FEB-1998; 98WO-FR00270.  
 XX  
 PR 28-JUL-1997; 97FR-0009587.  
 PR 11-FEB-1997; 97FR-0001574.  
 XX  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Bassilana F, Champigny G, Heutaux C, Lazdunski M;  
 PI Waldmann R, Lingueglia E;  
 XX  
 DR WPI: 1998-447231/38.  
 DR P-PSDB: AAM68507.  
 XX  
 PT Protein comprising proton-sensitive neuronal channel - useful for  
 PT screening for analgesics and for treating neurodegeneration  
 XX  
 PS Claim 15: Page 34-37; 64pp; French.  
 XX  
 CC This sequence represents the gene encoding a rat neuronal cationic  
 CC channel that is sensitive to amiloride and activated by protons and  
 CC is designated Acid Sensing Ionic Channel (ASIC). This sequence  
 CC represents the rat ASIC1B gene. The protein can be used to screen for  
 CC modulators of these channels, particularly to identify compounds that  
 CC modulate perception of acidity, as regards nociception (pain) rather than  
 CC taste. These compounds are used to treat or prevent pain associated with  
 CC acidity (e.g. in cases of inflammation, ischaemia or some tumours) and as  
 CC inhibitors of neurodegeneration caused by overexpression of the channels.  
 CC Antibodies to the protein are used to detect the channels in tissues, and  
 CC to act therapeutically as channel modulators. The nucleic acid can be  
 CC used to generate transgenic, particularly knockout, animals for studying  
 CC ASIC-related disorders, also for gene therapy. The channel protein,  
 CC or its (anti)agonists, can be used to treat or prevent cerebral  
 CC neurodegenerative conditions (e.g. Alzheimer's, Parkinson's or  
 CC Huntington's diseases, amyotrophic lateral sclerosis or cerebellar  
 CC ataxia).  
 XX  
 SQ Sequence 3647 BP; 895 A; 853 C; 900 G; 999 T; 0 other;

Query Match 80.0%; Score 16.8; DB 19; Length 3647;  
 Best Local Similarity 90.0%; Pred. No. 4.1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gctggcgaggttg99ggcc 20  
 ||||| |  
 DB 1863 GCTGGGGCAGTTGGGGGCC 1844

RESULT 11  
 AA261200/c  
 ID AA261200 standard; CDNA: 3647 BP.  
 XX

AC AA261200;  
 XX  
 DT 30-MAY-2000 (first entry)  
 XX  
 DE CDNA encoding a rat acid-sensitive cationic channel 1B (rASIC1B).  
 XX  
 DE Neuron acid-sensitive cation channel; ASIC; ASIC 1B;  
 KW proton-gated cation channel; biphasic desensitisation; amiloride;  
 KW cation transport channel; acid sensor; pH detection; ds.  
 XX  
 OS Rattus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 109..1788  
 FT /tag= a  
 FT /product= "acid-sensitive cationic channel 1B"  
 XX  
 PN W0200008149-A2.  
 XX  
 PD 17-FEB-2000.  
 XX  
 PE 05-AUG-1999; 99WO-IB01445.  
 XX  
 PR 05-AUG-1998; 98US-0095408.  
 XX  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Waldmann R, Bassilana F, Lazdunski M, De Welle JR;  
 PI WPI: 2000-195574/17.  
 DR P-PSDB: AAV69178.  
 XX  
 PT Novel human cation transport protein, Acid Sensing Ionic Channel 3 used  
 PT to identify substances capable of modulating cation transport channel  
 PT activity  
 XX  
 PS Disclosure; Page 73-76; 84pp; English.  
 XX  
 CC The present sequence encodes a rat neuronal acid-sensitive cation  
 CC channel 1B (ASIC1B) protein. The protein is a proton-gated cation  
 CC channel subunit that has biphasic desensitisation kinetics with both  
 CC a rapidly inactivating sodium-selective and a sustained component. The  
 CC channels are sensitive to amiloride. The specification describes ASIC3  
 CC proteins, which are expressed in the sensory neurons but not in the  
 CC brain. The cation transport channel proteins can be used in methods to  
 CC identify substances capable of modulating the activity of cation  
 CC transport channels. The human ASIC3 protein is also an acid sensor,  
 CC and might play an important role in the detection of lasting pH changes  
 CC in humans.  
 XX  
 SQ Sequence 3647 BP; 895 A; 853 C; 900 G; 999 T; 0 other;

Query Match 80.0%; Score 16.8; DB 21; Length 3647;  
 Best Local Similarity 90.0%; Pred. No. 4.1e+02;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gctggcgaggttg99ggcc 20  
 ||||| |  
 DB 1863 GCTGGGGCAGTTGGGGGCC 1844

RESULT 12  
 AAC75317  
 ID AAC75317 standard; CDNA: 4121 BP.  
 XX  
 AC AAC75317;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF872 polynucleotide sequence SEQ ID NO:1743.  
 DE Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;  
 KW

KM vulnerable; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
KM anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KM hypotensive; dermatological; immunosuppressive; antiinflammatory;  
KM antiviral; antibacterial; antifungal; antineumatic; antihyroid;  
KM antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
KM neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KM cholesterol ester storage; systemic lupus erythematosus; infection;  
KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KM allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KM bone damage; cartilage damage; nocturnal haemoglobinuria; burn; wound;  
KM thrombosis; contraceptive; ss.  
XX Homo sapiens.  
OS  
XX  
XX WO200058473-A2.  
XX  
XX 05-OC7-2000.  
XX  
XX  
XX 31-MAR-2000; 2000MO-US08621.  
XX  
XX 31-MAR-1999; 99US-0127607.  
XX 02-APR-1999; 99US-0127636.  
XX 05-APR-1999; 99US-0127728.  
XX 30-MAR-2000; 2000US-0540763.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Shinkets RA, Leach M;  
XX  
XX WPI: 2000-602362/57.  
XX P-PSDB: AABA1108.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
XX useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease -  
XX  
XX  
XX Claim 5; Page 1376-1378; 5507pp; English.  
XX  
XX AAC74446 to AAC77606 encode the proteins given in ABA40237 to ABA43397,  
XX which represent the human ORF open reading frames 1 to 3161. The ORF  
XX sequences have activities such as: cytostatic; hepatotropic; vulnerable;  
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
XX antidiabetic; hypotensive; dermatological; immunosuppressive;  
XX antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;  
XX antihyroid; and antianaemic. The sequences can be used for determining  
XX the presence of or predisposition to, or preventing or treating  
XX pathological conditions associated with an ORF-associated disorder. The  
XX nucleic acids can be used to express ORF proteins in gene therapy  
XX vectors. The proteins and nucleic acids may be used to treat cancers,  
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,  
XX graft vs host disease, cardiovascular disease, diabetes mellitus,  
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
XX bacterial and fungal infection, malaria, autoimmune disorders, asthma,  
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
XX nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
XX coagulation; to inhibit thrombosis; and as a contraceptive.  
XX  
XX Sequence 4121 BP; 818 A; 1186 C; 1260 G; 857 T; 0 other;

Query Match 80.0%; Score 16.8; DB 21; Length 4121;  
Best Local Similarity 90.0%; Pred. No. 4.1e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ctggcgaggtctggggcct 21  
||| ||||| ||||| |||||  
Db 2519 ctgagcgaggtctggggcct 2538

RESULT 13  
AAS28620  
ID AAS28620 standard; DNA: 14654 BP.  
XX  
XX AAS28620:  
AC  
XX  
XX 07-NOV-2001 (first entry)  
DT  
XX  
XX  
DE Genomic sequence #460 encoding for novel human respiratory antigen.  
XX  
XX Human; respiratory antigen; respiratory disorder; throat disorder;  
XX lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;  
XX anti allergic; anti asthmatic; anti inflammatory; olfactory;  
XX respiratory active; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200155448-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX  
XX 17-JAN-2001; 2001WO-US01333.  
XX  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 14-AUG-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225265.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226868.  
XX 23-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228294.  
XX 01-SEP-2000; 2000US-0229287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 01-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 05-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0231242.  
XX 08-SEP-2000; 2000US-0231243.  
XX 08-SEP-2000; 2000US-0231244.  
XX 08-SEP-2000; 2000US-0231413.  
XX 08-SEP-2000; 2000US-0231414.







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2002, 16:04:32 ; Search time 1959.41 Seconds  
(without alignments)  
224.280 Million cell updates/sec

Title: US-09-913-325-12

Perfect score: 21  
Sequence: 1 gctggcgagagctgggggacct 21

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hlg: \*  
3: gb\_in: \*  
4: gb\_cm: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_com: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pal: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_hlg\_hum: \*  
31: em\_hlg\_inv: \*  
32: em\_hlg\_other: \*  
33: em\_hlgo\_inv: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
------------	-------	--------------------	-------	-------------

C	1	21	100.0	975	6	AX014321
C	2	21	100.0	1416	6	HUMSGLY
C	3	21	100.0	1512	6	AX302541
C	4	21	100.0	1512	9	HUMABOJ
C	5	21	100.0	1648	9	HUMTRPM2A
C	6	21	100.0	1658	9	BC019588
C	7	21	100.0	1676	6	HSCSP40
C	8	21	100.0	1678	6	AX202086
C	9	21	100.0	1684	6	BC010514
C	10	21	100.0	3196	6	AX330207
C	11	21	100.0	3196	6	AX337122
C	12	21	100.0	3196	6	HUMTRPM2A4
C	13	21	100.0	8133	9	HUMSEP40A
C	14	21	100.0	187453	2	AC025192
C	15	21	100.0	234431	9	AF311103
C	16	18.4	87.6	57722	2	HSCB3B10
C	17	18.4	87.6	76295	2	AC023542
C	18	18.4	87.6	86501	9	AC079171
C	19	18.4	87.6	106008	9	HS506
C	20	17.8	84.8	398	12	SYNSV4PYRO
C	21	17.8	84.8	2132	9	HSAB01802
C	22	17.8	84.8	73396	9	AC008848
C	23	17.8	84.8	114653	2	AC092411
C	24	17.8	84.8	143804	2	AC012617
C	25	17.8	84.8	157758	9	AC098824
C	26	17.8	84.8	169638	9	AC009433
C	27	17.8	84.8	178000	2	AC011847
C	28	17.8	84.8	185854	2	AC027736
C	29	17.8	84.8	187640	2	AC095241
C	30	17.8	84.8	194451	2	AC090647
C	31	17.4	82.9	119452	2	AC105186
C	32	17.4	82.9	140453	2	AC095285
C	33	17.4	82.9	183584	2	AC095195
C	34	17.4	82.9	216425	2	AC064825
C	35	17.4	82.9	220376	2	AC025579
C	36	17.4	82.9	241714	2	AC083834
C	37	17	81.0	9738	14	AF164485
C	38	16.8	80.0	724	9	HSAB30907
C	39	16.8	80.0	1111	4	PICALIP1
C	40	16.8	80.0	1242	4	AF177912
C	41	16.8	80.0	1639	4	BTU08213
C	42	16.8	80.0	2170	9	AB03799501
C	43	16.8	80.0	2535	6	AX179295
C	44	16.8	80.0	2962	6	A98491
C	45	16.8	80.0	2962	10	RNC6519

## ALIGNMENTS

RESULT	1	AX014321	975 bp	DNA	linear	PAT 07-SEP-2000
LOCUS		AX014321/C				
DEFINITION		Sequence 29 from Patent WO954353.				
ACCESSION		AX014321				
VERSION		AX014321.1 GI:10040675				
KEYWORDS						
SOURCE						
ORGANISM		human.				

human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
1 (bases 1 to 975)  
Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and  
Pilarczyk,C.  
Human nucleic acid sequences of normal uterus tissue  
Patent: WO 954353-A 28 OCT-1999.

## TITLE

JOURNAL  
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN  
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN  
(DE); PILARSKY CHRISTIAN (DE)

## FEATURES

source	1..975
	/organism="Homo sapiens"
	/db_xref="taxon:9606"

BASE COUNT	221 a	320 c	248 g	186 t
ORIGIN				
Query Match	100.0%;	Score 21;	DB 6;	Length 975;
Best Local Similarity	100.0%;	Pred. No. 26;		
Matches 21:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	1 gctgggcgagcttg9ggagcct 21			
Db	903 GCTGGCGGAGTGTGGGGGCGCT 883			
RESULT 2				
HUMSGLY/c		1416 bp	RNA	linear
LOCUS	HUMSGLY			PRI 09-JAN-1995
DEFINITION	Human sulfated glycoprotein-2 mRNA, 3'end.			
ACCESSION	M74816			
VERSION	M74816.1			
KEYWORDS	GI:338056			
SOURCE	secreted protein, sulfated glycoprotein-2.			
ORGANISM	Homo sapiens male 63 yr. old adult brain tumor cDNA to mRNA.			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	Danik,M., Chabot,J.G., Mercier,C., Benabid,A.L., Chauvin,C., Ourlon,R. and Suh.M.			
JOURNAL MEDLINE	Human gliomas and epileptic foci express high levels of a mRNA related to rat testicular sulfated glycoprotein 2, a purported marker of cell death			
FEATURES	Proc. Natl. Acad. Sci. U.S.A. 88 (19), 8577-8581 (1991)			
source	92020896			
	Location/Qualifiers			
	1. 1416			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/map="8"			
	/sex="male"			
	/cell_type="astrocytoma (gr. III)"			
	/tissue_type="brain tumor"			
	/dev_stage="63 yr. old adult"			
gene	1. 1416			
	/gene="CLT"			
	<1. 1172			
CDS	/gene="CLT"			
	/codon_start=3			
	/product="sulfated glycoprotein-2"			
	/protein_id="AA60321.1"			
	/db_xref="GI:338057"			
	/db_xref="GDB:600-125-226"			
	/translation="EKTNEERKTLTSLNEARKKKEDALNETRESETKLRELPGVNEB			
	TYMALWEBCPKLQKOTCMKFAVRCGSGGLVQGLEFLNOSSPFYFWNGDRIDSL			
	LENDROTHMLDVMQDHFSRASIIIDELFODRFTREPODYHYLPSLRPRHFFEF			
	PKSIVTSLMPEFSPEYELPNFAMPOPLEMHEHQAMDIFHSAPQHPETEIRGGEG			
	DDDTGAREITHNYSKMLNTSGCLRMKDDCKREITLSVCSNNPSQAKRLREDSLOVAR			
	LTGRYNELKISYCKMKLNTSGCLRMKDDCKREITLSVCSNNPSQAKRLREDSLOVAR			
	DSDPVSGTEYVVALPEPDPITTYVPPVPEVSKNPKFMETAKRLQETRKRRKEE"			
BASE COUNT	344 a	440 c	365 g	267 t
ORIGIN				
Query Match	100.0%;	Score 21;	DB 9;	Length 1416;
Best Local Similarity	100.0%;	Pred. No. 25;		
Matches 21:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	1 gctgggcgagcttg9ggagcct 21			
Db	1309 GCTGGCGGAGTGTGGGGGCGCT 1289			
RESULT 3				
AX302541/c				
LOCUS	AX302541	1512 bp	DNA	linear
				PAT 30-NOV-2001

DEFINITION	Sequence 59 from Patent WO015177.
AX302541	
VERSION	AX302541.1 GI:17383080
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE	1 (sites)
JOURNAL	Morin,P.J., Sherman-Bauslt,C.A., Pizer,E.S. and Hough,C.D. Tumor markers in ovarian cancer Patent: WO 0175177-A 59 11-OCT-2001; THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES	Location/Qualifiers
source	1. .1512 /organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	378 a 461 c 387 g 286 t
ORIGIN	
Query Match	100.0% Score 21; DB 6; Length 1512;
Best Local Similarity	100.0%; Prcd. No. 24;
Matches 21: Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 1	gctggcgcgagttggggcct 21 
Db 1405	gctggcgcgagttggggcct 1385
RESULT 4	
HUMAPOJ/c	1512 bp mRNA linear PRI 31-OCT-1994
LOCUS	HUMAPOJ
DEFINITION	Human apolipoprotein J mRNA, complete cds.
ACCESSION	J02908
VERSION	J02908.1 GI:178854
KEYWORDS	apolipoprotein J; high density lipoprotein.
SOURCE	Human fetal liver, cDNA to mRNA, clone lambda[1-3].
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE	1 (bases 1 to 1512)
JOURNAL	de Silva,H.V., Harmony,J.A., Stuart,W.D., Gil,C.M. and Robbins,J.
MEDLINE	Apolipoprotein J: structure and tissue distribution Biochemistry 29 (22), 5380-5389 (1990) 90344779
COMMENT	Draft entry and computer-readable sequence for [1] kindly submitted by J.Robbins, 21-MAR-1990.
FEATURES	Location/Qualifiers
source	1. .1512 /organism="Homo sapiens" /db_xref="taxon:9606" /map="Unassigned" <1. .1512 /gene="APOJ" /product="apolipoprotein J mRNA" 1. .1512 /gene="APOJ" 18. .1268 /note="apolipoprotein J precursor" /codon_start=1 /protein_id="AAAS1765.1" /db_xref="GI:178855" /db_xref="GDB:13053" /translation="WSNKGKRYNKELEQNAVNGVQKIKTLKTNKEKTLTSLDEEA KKKKDALNRESEETKLKELPGVCNFTMAAMEECKPCKQTCMKFYARCRSGSL VORRLEELFINSPEFYFMNGDRIDSLDENRQOTIMLDVMDHFSRASSIIDLFODD RFTFEPODPTVHLPESLPHRPHFPFKSRIVRSIMPSPEPLNFAMQOPLEMI HQAQOQIDIFHSPAFQPHPTFRREGDDRTVCRIRHNSGCLRMQDQCKRETLI SDVCSGNNSQAKLRRELDSEIQVAERLPRKVELIKSYQMKMLTSSLELDQENR KVSRLANTQGEQDQYILRIYTVASHTSDSPVSGTYEYVVKLFDSDDITTVIVPEVSR KAPKPMETVAKALQETRKHRE"





ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 1658;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctggcgagtggtgggacct 21  
|||||  
Db 1535 GCTGGCGGAGTGTGGGGCCT 1515

RESULT 7  
HSCSP40 1676 bp mRNA linear PRI 22-MAR-1995  
LOCUS Human SP-40, 40 mRNA for complement-associated protein SP-40, 40  
DEFINITION alpha-1 and beta-1 chain.  
ACCESSION X14723.1 GI:30250  
VERSION X14723.1  
KEYWORDS complement-associated protein; serum protein; SP-40,40 gene.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE 1 (bases 1 to 1676)  
JOURNAL Kirszbaum, L.  
Direct Submission  
Submitted (17-MAR-1989) Kirszbaum L., The University of Melbourne,  
The Precinical Centre, School of Veterinary Science, Parkville  
Victoria 3052, Australia  
2 (bases 1 to 1676)  
Kirszbaum, L., Sharpe, J.A., Murphy, B., d'Apice, A.J., Classon, B.,  
Hudson, P. and Walker, I.D.  
REFERENCE Molecular cloning and characterization of the novel, human  
TITLE complement and reproductiv systems  
JOURNAL EMBO J. 8 (3), 711-718 (1989)  
COMMENT The sequence overlaps with that reported by Murphy et. al. in J.  
Clin. Invest. 81:1858-1864(1988).  
FEATURES  
source location/Qualifiers  
1..1676  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="LK (107)"  
/issue\_type="liver"  
/clone\_lib="lambda gtl1"  
48..113  
48..1397  
/note="SP-40,40 prepropeptide (AA -22 to 427)"  
/codon\_start=1  
/protein\_id="CAA32847.1"  
/db\_xref="GI:30251"  
/db\_xref="SWISS-PROT:P10909"  
/translation="MMKTLFLFGLLITWSSGOVIGDQTSNDLEQMSNGSKYVNK  
ETONAVNGVQKITLLEKTEERKTLSTLEAKKKKEDLNETRESERFLAKLPVC  
NLEMMALMECKPCDLKQCKFYARVCRSGSLVGRLEFLNQSSPEFYKWMNGDRD  
SILENDRQCTHMLDMQDHFSSASIIDELFORFREPQDTHYLPFLPRRHF  
FFPKSRIVRSIPYEPPLNFHMFQPLFEMIHQAQAMDIHFHSPAFQHPPEER  
EGDDRTVCREIRHNSGCLRMKDCCKREIISVDCSTNNPSQAKLRRLDSLOVA  
ERLTKRNELLSKQMKMLNTSSLEBNQFQFVWSLALTLGGEQGYLRRTVASH  
TSDSDVPSGVTEVYVKLFSDSPITVIVPEVSRKNPMEFVAKALQETRRKRRP"  
114..>114  
/note="beta-chain"  
114..1394  
/note="SP-40,40 prepeptide (AA 1-427)"  
729..1394  
/product="mature alpha-chain (AA 205-427)"  
1622..1627  
/note="pot. polyA signal"  
BASE COUNT 436 a 488 c 437 g 315 t  
ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 1676;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctggcgagtggtgggacct 21  
|||||  
Db 1536 GCTGGCGGAGTGTGGGGCCT 1516

RESULT 8  
AX202086/c 1678 bp DNA linear PAT 30-AUG-2001  
LOCUS Sequence 16 from Patent WO0153531.  
ACCESSION AX202086  
VERSION AX202086.1 GI:15391872  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE 1 (bases 1 to 1678)  
JOURNAL Phippard, D., Vasanthakumari, G., Dolson, S. and Ma, X.J.  
Osteoarthritis tissue derived nucleic acids, polypeptides, vectors,  
and cells  
Patent: WO 0153531-A 16 26-JUL-2001;  
Pharmacia Corporation (US)  
FEATURES  
source location/Qualifiers  
1..1678  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 424 a 491 c 450 g 313 t  
ORIGIN

Query Match 100.0%; Score 21; DB 6; Length 1678;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctggcgagtggtgggacct 21  
|||||  
Db 1558 GCTGGCGGAGTGTGGGGCCT 1538

RESULT 9  
BC010514/c 1684 bp mRNA linear PRI 12-JUL-2001  
LOCUS Homo sapiens, clusterin (complement lysis inhibitor, SP-40,40,  
DEFINITION sulfated glycoprotein 2, testosterone-repressed prostate message 2,  
apolipoprotein J), clone MGC:18080 IMAGE:4150452, mRNA, complete  
cds.  
ACCESSION BC010514  
VERSION BC010514.1 GI:14714740  
KEYWORDS MGC.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE 1 (bases 1 to 1684)  
JOURNAL Strausberg, R.  
Direct Submission  
Submitted (10-JUL-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIR-MGC Project URL: http://mgc.ncl.nih.gov  
Contact: MGC help desk  
Email: cgaps-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
DNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center

Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [villalobebcm.tmc.edu](mailto:villalobebcm.tmc.edu)  
Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRK Plate: 19 Row: 0 Column: 12.

#### FEATURES

Location/Qualifiers

1..1684  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:18080 IMAGE:4150452"  
/issue\_type="Brain, glioblastoma with EGFR amplification"  
/clone\_id="NCI\_CGAP\_Brn64"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"  
69..1418  
/codon\_start=1  
/product="clusterin (complement lysis inhibitor, SP-40/40, sulfated glycoprotein 2, testosterone-repressed prostate message 2, apolipoprotein j)"  
/protein\_id="AAH10514.1"  
/db\_xref="GI:14714741"  
/translation="MMKTLIFVGLITWESGOVIGDPTVSNETLOEPMNOSKYYNK EIONAVNGVQITLIEKTEKREKTLISNLEBKKEKELNLTRESEKLEDELGVC NETMALWECRCKPCLKOTCKEFARVCRSGSLVGQLEPLNOSPFYFWMNGDID SLENDROQTHMLDVMDHFRASSIIDELEFDRFTRPDPTVHYLPSPHRRHF FPEKSRVRSIMPFSPYEPLEHAMEQPLEMEHQAQAMHIFHSPAQHPTEIR EGGDDRTVCREIRHNSGCLRMKODCKCEILSYCCNNPSOAKLRELESLQVH ERLTRYNELKSYOMKMLNTSLLEIOLNDFWMLANLNOGEOVYLYRTVYVSH TSDSDVPSGTEVYVLFDSPTITVTPVPSVSKNPKFMEYAEKALQYRKRRHE"

#### CDS

#### BASE COUNT

433 a 493 c 448 g 310 t

#### ORIGIN

#### Query Match

Best Local Similarity 100.0%; Score 21; DB 9; Length 1684;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctggcgagtggtgggacct 21

|||||

Db 1555 GCTGGCGGAGTGGGGCCT 1535

#### RESULT 10

#### AX330207/c

#### LOCUS

#### DEFINITION

#### ACCESSION

#### VERSION

#### KEYWORDS

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### FEATURES

#### source

#### BASE COUNT

#### ORIGIN

AX330207 3196 bp DNA linear PAT 09-JAN-2002  
Sequence 716 from Patent WO0194629.  
AX330207  
AX330207.1 GI:18103185  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (sites)  
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrigan, S., Soppet, D.R. and Weaver, Z.  
Cancer gene determination and therapeutic screening using signature  
gene sets  
Patent: WO 0194629-A 716 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
Location/Qualifiers  
1..3196  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

833 a 762 c 789 g 812 t

Query Match 100.0%; Score 21; DB 6; Length 3196;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctggcgagtggtgggacct 21

|||||

Db 2891 GCTGGCGGAGTGGGGCCT 2871

#### RESULT 11

#### AX337122/c

#### LOCUS

#### DEFINITION

#### ACCESSION

#### VERSION

#### KEYWORDS

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### FEATURES

#### source

AX337122 3196 bp DNA linear PAT 09-JAN-2002  
Sequence 7631 from Patent WO0194629.  
AX337122  
AX337122.1 GI:18127841  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (sites)  
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrigan, S., Soppet, D.R. and Weaver, Z.  
Cancer gene determination and therapeutic screening using signature  
gene sets  
Patent: WO 0194629-A 7631 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
Location/Qualifiers  
1..3196  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

#### BASE COUNT

833 a 762 c 789 g 812 t

#### ORIGIN

#### Query Match

Best Local Similarity 100.0%; Score 21; DB 6; Length 3196;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctggcgagtggtgggacct 21

|||||

Db 2891 GCTGGCGGAGTGGGGCCT 2871

#### RESULT 12

#### HUMTRPM24/c

#### LOCUS

#### DEFINITION

#### ACCESSION

#### VERSION

#### KEYWORDS

#### SEGMENT

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### MEDLINE

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### FEATURES

#### source

HUMTRPM24 3196 bp DNA linear PRI 23-AUG-1996  
Human TRPM-2 protein gene, exons 7, 8, 9 and complete cds.  
M63379.1 GI:292841  
TRPM-2 protein.  
4 of 4  
Homo sapiens DNA.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (sites)  
Wong, P., Pineault, J., Lakins, J., Taillefer, D., Leger, J., Wang, C.  
and Tenniswood, M.  
Genomic organization and expression of the rat TRPM-2 (clusterin)  
J. Biol. Chem. 268 (7), 5021-5031 (1993)  
93186813  
2 (bases 1 to 3196)  
Wong, P., Taillefer, D., Lakins, J., Pineault, J., Chader, G. and  
Tenniswood, M.  
Molecular characterization of human TRPM-2/clusterin, a gene  
associated with sperm maturation, apoptosis and neurodegeneration  
Eur. J. Biochem. 221 (3), 917-925 (1994)  
94237156  
Location/Qualifiers  
1..3196

#### BASE COUNT

833 a 762 c 789 g 812 t

#### ORIGIN

/organism="Homo sapiens"
/db\_xref="taxon:9606"
join(M63376.1:1422..1469,M63376.1:5509..5634,
M63376.1:7021..7169,M63377.1:155..325,M63378.1:308..719,
M63378.1:1247..1351,995..1224,2362..2537,2745..2998)
/gene="TRPM-2"
join(M63376.1:1422..7610,M63377.1:1..940,M63378.1:1..1634,
1..3166)
/gene="TRPM-2"
join(M63376.1:5538..5634,M63376.1:7021..7169,
M63377.1:155..325,M63378.1:308..719,M63378.1:1247..1351,
995..1224,2362..2537,2745..2754)
/gene="TRPM-2"
/codon\_start=1
/protein\_id="AAB06507.1"
/db\_xref="GI:292843"
/translation="MMKTLIFVGLITLWESGOVLGDTVSDNLEOEMSGSKYVK
ETQANVGVKQITLIEKTNDEKTLISLEBKAKKEDALNETRESEYKLKELPGVC
NETMMALMECKPCQKQYKAVRCSSGLVGRQLEELNOSSEFYEMNGDRID
SLENDROQTHMLDMQDHFSSASSIIDELFODREPDYHYLPESLPHRRHF
EFPKSRIVRSIAMPSPYEPFLNFMQPLFMIEHAQOAMDIHFHSPAFQHPTEFR
EGDDRTVCREIRHNSGCLRMKDQCKREITISVDCSTNNSOAKRLREDSLOVA
ERLTRKNELLSQYQKMLNFSLEQINQOPFWYSRLATLQGEDOYLRVTVASH
TSDDVPSGVTEVVKLFDSDPITVIVPEVSRKRFMETVAKALOETRKAREE"
995..1224
/gene="TRPM-2"
/number=7
2362..2537
/gene="TRPM-2"
/number=8
2745..2998
/gene="TRPM-2"
/number=9
3085..3090
/gene="TRPM-2"
3146..3151
/gene="TRPM-2"
3161..3166
/gene="TRPM-2"
BASE COUNT 833 a 762 c 789 g 812 t
ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 3196;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gctggcgaggttggggacct 21
|||||
Db 2891 GCTGGCGGAGTTGGGGCCT 2871

RESULT 13
HUMSP4040A/c 8133 bp DNA linear PRI 13-JAN-1995
LOCUS HUMSP4040A 8133 bp DNA linear PRI 13-JAN-1995
DEFINITION Homo sapiens SP40.40 gene, exons 5-9.
ACCESSION L00974
VERSION L00974.1 GI:338304
KEYWORDS
SOURCE Homo sapiens Foetal Liver DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 8133)
AUTHORS Giew.M.D., Kirsbaum.L., Bozas,S.E. and Melker,I.D.
TITLE Partial nucleotide sequence of the human SP40.40 gene
JOURNAL Unpublished (1993)
FEATURES
Location/Qualifiers
1..8133
/organism="Homo sapiens"
/db\_xref="taxon:9606"
/tissue\_type="Foetal Liver"
520..7869
gene

CDS
/gene="SP40.40"
join(520..931,1460..1564,5852..6081,7221..7396,7605..7614)
/partial
/gene="SP40.40"
/codon\_start=1
/product="SP40.40"
/protein\_id="AAB0567.1"
/db\_xref="GI:338305"
/translation="LEEFINQSSPEYFMWNGDRIDSLLENDROQTHMLDMQDHFSSRA
SSIDELFODREPDYHYLPESLPHRRHFEPKSRIVRSIAMPSPYEPFLNFH
AMOPFLMIEHAQOAMDIHFHSPAFQHPTEFRSGDDRTVCREIRHNSGCLRMK
DQCKDREITLSDCSTNNSOAKRLREDESIQVAERLTRKNELLSQYQKMLNFS
LEQLNEQFNWVSRILNTQGEDOYLRVTVASHTSDDVPSGVTEVVKLFDSDP
ITVIVPEVSRKRFMETVAKALOETRKAREE"
520..931
/gene="SP40.40"
/number=5
1460..1564
/gene="SP40.40"
/number=6
5852..6081
/gene="SP40.40"
/number=7
7221..7396
/gene="SP40.40"
/number=8
7605..7869
/gene="SP40.40"
/number=9
7838..7843
/gene="SP40.40"
8008..8013
POLYA\_signal
8023..8028
BASE COUNT 1982 a 2024 c 2002 g 2125 t
ORIGIN

Query Match 100.0%; Score 21; DB 9; Length 8133;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gctggcgaggttggggacct 21
|||||
Db 7752 GCTGGCGGAGTTGGGGCCT 7732

RESULT 14
AC025192/c 187453 bp DNA linear HTG 13-MAY-2001
LOCUS AC025192 187453 bp DNA linear HTG 13-MAY-2001
DEFINITION Homo sapiens chromosome 8 clone RP11-509E2 map 8, WORKING DRAFT
SEQUENCE, 7 unordered pieces.
ACCESSION AC025192
VERSION AC025192.3 GI:12313839
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 187453)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Batra,N., Bastien,V., Bede,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choedel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeAtellano,K., Dewar,K., Diaz,J.S.,
Dodgson,S., Domingo,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heathford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Lakocque,K., Lamazares,R., Landers,T., Lehoczy,J.,

```

TITLE
JOURNAL
COMMENT
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L5592
Center clone name: 509_E2
----- Summary Statistics
Sequencing vector: M13; M77815; 48% of reads
Sequencing vector: Plasmid; n/a; 52% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 185120 bases at least Q40
Consensus quality: 186074 bases at least Q30
Consensus quality: 186533 bases at least Q20
Insert size: 175000; agarose-fp
Insert size: 186853; sum-of-contigs
Quality coverage: 8.9 in Q20 bases; agarose-fp
Quality coverage: 8.3 in Q20 Da.
NOTE: This is a 'working draft' sequence. It currently
consists of 7 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 17773: contig of 17773 bp in length
17774 17873: gap of 100 bp
17874 20278: contig of 2405 bp in length
20279 20378: gap of 100 bp
20379 35742: contig of 15364 bp in length
35743 35842: gap of 100 bp
35843 65413: contig of 29571 bp in length
65414 65513: gap of 100 bp
65514 112793: contig of 47280 bp in length
112794 112893: gap of 100 bp
112894 150031: contig of 37138 bp in length
150032 150131: gap of 100 bp
150132 187453: contig of 37322 bp in length.
location/Qualifiers
1 .187453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-509E2"
/clone_lib="RPCI-11 Human Male BAC"
1 .17773
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
17874..20278
/note="assembly_firagment"
misc_feature
20379..35742
/note="assembly_fragment"
misc_feature
35843..65413
/note="assembly_fragment"
misc_feature
65514..112793
/note="assembly_fragment"
misc_feature
112894..150031
/note="assembly_fragment"
misc_feature
150132..187453
/note="assembly_fragment"
clone_end:r7
vector_side:right"
BASE COUNT 51878 a 43372 c 44038 g 47545 t 600 others
ORIGIN
Query Match 100.0%; Score 21; DB 2; Length 187453;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 gcctggcgcgagctggcgccct 21
Db 137322 GTTGGGCGAGTTGGGGCCT 137302
RESULT 15
AF311103/c 234431 bp DNA linear PRI 06-FEB-2002
LOCUS Homo sapiens chromosome 8 clone SCB-212e3 map 8p12, complete
DEFINITION sequence.
ACCESSION AF311103 GI:18542957
VERSION AF311103.3 HTG: HTGS_FULFLOP; HTGS_ACTIVEFIN.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 234431)
Wen,G., Bleischmidt,K., Baumgart,C., Taudin,S., Baumgart,C. and
Platzter,M.
Chromosome 8 genomic sequence
Unpublished
2 (bases 1 to 234431)
Bleichschmidt,K., Jandrig,B., Baumgart,C., Dette,M.D., Jahn,N.,
Wenzel,U., Schilhelbel,M.B., Wen,G., Tauden,S. and Rosenthal,A.
Direct Submission
Submitted (04-OCT-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
3 (bases 1 to 234431)
Wen,G. and Platzter,M.
Direct Submission
Submitted (06-FEB-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
On Feb 6, 2002 this sequence version replaced gi:14389409.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: H319
Center clone name: SCB-212e3
----- Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 233404 bases at least Q40
Consensus quality: 234413 bases at least Q30
Consensus quality: 234428 bases at least Q20
Quality coverage: 5.81 x in Q20 bases; sum-of-contigs
Sequence Quality Assessment:
This entry has been annotated with sequence quality

```

estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
GenBank flat file format but are available as part  
of this entry's ASN.1 file.

FEATURES  
Source

Location/Qualifiers  
1. .234431  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p12"  
/clone="SCb-212e3"

BASE COUNT 62993 a 56257 c 55890 g 59291 t  
ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 21; DB 9; Length 234431;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gctgggcyagltggggcct 21  
|||||  
DB 26443 GCTGGCGAGTTGGGGCCT 26423

Search completed: May 17, 2002, 16:04:42  
Job time: 16142 sec